Characterization of β 1,3 Glycosyltransferases from *Drosophila melanogaster*

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde (Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Reto Müller

von

Grüsch GR

Promotionskomitee

Prof. Dr. Thierry Hennet (Vorsitz)

Prof. Dr. Monika Steinmann-Zwicky

Prof. Dr. Hans Rudolf Bosshard

Zürich 2005

Acknowledgements

Science is team work. Therefore, there are a lot of people to be thanked at the end of a Ph.D. First, I would like to thank **Prof. Dr. T. Hennet**, for his supervision and for his introducing me to the fascinating world of Glycobiology. Both Biochemistry and Genetics had to be looked at in these projects. **Prof. Dr. H. R. Bosshard** from the Institute of Biochemistry and **Prof. Dr. M. Steinmann-Zwicky** from the Institute of Zoology kindly shared the responsibility for my entry into the scientific world as my "doctor parents".
I am grateful for invaluable technical and methodological advice from Drs. **Erich Frei** (from the group of **Prof. Dr. M. Noll**, at the IMB, Zurich), **Albana Rexhepaj** (Physiology, Zurich), **Tim Henion** (Boston), **Dapeng Zhou** (Chicago)**, Andreas Hülsmeier** (Physiology), **Jack Rohrer** (Physiology) and **Prof. Dr. E. Berger** (Physiology). **Michi Daube** and **Werner Boll** (IMB) introduced me to fly husbandry and oocyte injection. Collaborating with Professors **Raffi Aroian** (San Diego), **Fritz Altmann** (Vienna), **Mike Tiemeyer** (Athens) and **Kelly Ten Hagen** (Bethesda) was a pleasure.
Technical assistance from **Bea Berger** (occasional cell culturing) has saved me a man year and **Marianne Farah** helped subcloning and injecting dsRNA.
Claudia Grubenmann, **Jacqueline Stöckli** and **Franziska Biellmann** from the Physiology L-floor were fun to be with.

Arianna Bisaz has kept me almost sane and house sharing community partners and friends like **Alexander Spreiter**, **Lothar Birke**, **Jacques Lecoultre**, **Paul Wersin**, **Cecilia Hurley**, **Christoph Fuchs** and **Katrin Zeigerer** made my day when research did not and left me houseroom while writing this up.

My mother, **Edith Müller**, supported me in every way during my studies and during my Ph.D., thank you. This one is for you.

Thesis Summary and Aim

Glycosylation is the enzymatic modification of various compounds with sugars. In the present work, glycosylation of proteins and lipids is investigated. To be able to carry out glycosylation reactions, the cell relies on an enzymatic machinery to synthesize activated sugars and to transport these to the proper place within the cell. Then, enzymes transfer the sugar onto a specific acceptor substrate. These enzymes are called glycosyltransferases. The first glycosylation reaction may be followed by additional ones, elongating the initial unit to a sugar chain.

The aim of my thesis was to characterize members of a family of enzymes from *Drosophila melanogaster* that elongate the sugar chains of glycolipids and glycoproteins. Sequence homology between these and related enzymes suggested that they all form a β1,3 linkage. The fly was chosen as model system to have a basis to later genetically investigate glycosyltransferase mutants.

Mutations in one of the characterized enzymes, a β1,3 glycosyltransferase encoded by *brainiac*, are lethal and embryos display an enlarged central nervous system compared to the *wild-type* animal. Activity determination and product analysis showed, that the enzyme transfers the sugar *N*-acetylglucosamine onto a mannose in the β configuration. These donor and acceptor preferences suggested involvement of the enzyme in the synthesis of the arthro-series core structure of insect glycolipids, N-acetylglucosamine β1,3 mannose β1,4 glucose β-ceramide. Indeed, *in vitro* activity towards glycolipids could be established. The arthro-series glycolipid core is not found in vertebrates and attempts to complement the *brainiac* mutant by transgenic expression of mammalian enzymes displaying sequence similarities to Brainiac failed. However, we were able to correct the mutant phenotype with a sequence-similar *Caenorhabditis elegans* transgene, "*Bacillus thuringiensis* resistance-5" (bre-5). *Caenorhabditis elegans* mutants for this *brainiac*-orthologous gene have an only slightly reduced viability and fertility but escape the lethal effects of being fed a specific *bacillus thuringiensis* toxin. *Bacillus thuringiensis* toxins are extensively used as a means for pest control in agriculture and are also transgenically expressed in wheat, soy bean and corn. In the first place, the rescue of the *Drosophila melanogaster brainiac* mutant by the *Caenorhabditis elegans* enzyme shows that *Bacillus thuringiensis* toxins can bind glycolipids. The binding to the arthro-series core partially explains why animals that do not have this glycolipid-core, are not affected by this toxin. Secondly, the difference in resistance levels or in the acquisition of resistance to *Bacillus thuringiensis* toxins by different organisms may be explained by the (lack of) vital importance of the receptor-glycolipid for the organism,

since *Caenorhabditis elegans* evidently survives without this arthro-series core, while *Drosophila melanogaster* does not.

Investigation of β1,3 glycosyltransferases was then extended to a large family of *Drosophila melanogaster* genes comprising nine members. The proteins encoded by these genes displayed extensive sequence similarity to the mammalian β1,3 galactosyltransferase enzyme, which synthesizes the T-antigen. The T-antigen is the disaccharide galactose β1,3 N-acetylgalactosamine α-O-Serine/Threonine, which represents the core structure of mucin-protein O-glycosylation. The T-antigen is formed by most animals. While mammals harbour only one T-antigen- synthesizing enzyme, we found four enzymes in *Drosophila melanogaster*, which show the corresponding activity. Transcripts of the most active of these enzymes were found to be deposited by the mother into the embryo and later to be expressed in the amnioserosa, a transient tissue that vanishes during embryogenesis. The amnioserosa – epidermis interaction is considered to be a wound-healing model, because it separates the lateral and the dorsal epidermis during early stages of embryonic development, then disappears completely to allow the two tissues to coalesce. Transcripts of two other enzymes were found in the salivary glands. The developmental and physiological roles of these enzymes in *Drosophila melanogaster* await the characterization of mutant animals. However, the enzymological basis for the genetic analysis of the T-antigen was laid with the present work.

Zusammenfassung und Ziel

Unter Glykosylierung versteht man die enzymatische Derivatisierung von verschiedensten Verbindungen mit Zuckern. In dieser Arbeit ging es um die Glykosylierung von Proteinen und Lipiden. Um glykosylieren zu können, benötigt die Zelle Enyzme und Energie, um aktivierte Zucker herzustellen, und um diese an den richtigen Platz in der Zelle zu transportieren. Schliesslich braucht es Enzyme, welche die eigentlichen Glykosylierungsreaktionen durchführen. Diese Enzyme heissen Glykosyltransferasen. Dem ersten Glykosylierungsschritt folgt bisweilen eine Verlängerung zu einer Zuckerkette durch weitere Glykosyltransferasen.

Das Ziel meiner Dissertation war die Charakterisierung einer Familie von Glykosyltransferasen, deren Mitglieder die Zuckeranteile von Glykoproteinen und Glykolipiden mit anderen Zuckermolekülen in einer β1,3 Bindung verknüpfen. β1,3 Glykosyltransferasen weisen untereinander Ähnlichkeiten in ihrer Aminosäurensequenz auf, unabhängig davon, welchen Zucker sie auf welchen anderen Zucker transferieren. Um herauszufinden, ob und wie wichtig β1,3 verknüpfte Zuckerverbindungen sind, haben wir β1,3 Glykosyltransferasen von *Drosophila melanogaster* untersucht, damit später Mutanten auf genetische Interaktionen getestet werden können.

Für eines der untersuchten Enzyme gab es bereits Mutanten. Homozygote *brainiac* - Mutanten sterben und weisen als Embryonen ein abnormal vergrössertes zentrales Nervensystem auf. Bei der biochemischen Charakterisierung des Enzyms stellte sich heraus, dass die β1,3 Glykosyltransferase Brainiac den Zucker *N*-Acetylglukosamin auf Mannose in der β Konfiguration überträgt und dass es an der Herstellung von Arthropoden-Glykosphingolipiden beteiligt ist, indem es deren Kernstruktur, Mannose β1,4Glukose β-Ceramide, mit *N*-Acetylglukosamin elongiert. Tatsächlich konnten wir *in vitro* Aktivität von Brainiac gegen Insekten-Glykolipide nachweisen. Während der Versuch scheiterte, den *brainiac*-Phänotyp durch Transgene von sequenz-ähnlichen Säugetier-Enzymen zu korrigieren, gelang dies mit Hilfe eines *Caenorhabditis elegans* Transgens. Die *Caenorhabditis elegans* Mutanten in diesem *brainiac*-orthologen Gen, *bre-5* (*bacillus thuringiensis* resistance), überleben und sind fertil, hingegen sind sie resistent gegen ein *Bacillus thuringiensis* Toxin. Diese Toxine werden häufig zur Bekämpfung von Schadinsekten in der Landwirtschaft eingesetzt und aus demselben Grund auch transgen in Mais, Soja, Weizen und Tabak exprimiert. Die transgene Komplementation der lethalen *Drosophila melanogaster brainiac* Mutante durch das *Caenorhabditis elegans* Enzym beweist erstens, dass *Bacillus thuringiensis* Toxine (Bt-Toxine) Glykolipide binden können, die nur in Arthropoden vorkommen. Das erklärt, weshalb Vertebraten keine Schäden durch diese Bt-

Variante erleiden. Zweitens impliziert dieses Resultat, dass *Drosophila melanogaster* kaum dieselbe Resistenz gegen dieses Toxin entwickeln wird wie *Caenorhabditis elegans*, da das Bt-Toxin bindende Glykolipid für die Fliege überlebenswichtig ist.

Die Untersuchung der β1,3 Glykosyltransferasen wurde sodann auf eine Familie von β1,3 Glykosyltransferasen mit neun Mitgliedern ausgedehnt, welche Sequenzähnlichkeiten zu einem Säugetierenzym aufweist, welches das T-Antigen herstellt. Das T-Antigen ist das Disaccharid Galaktose β1,3 *N*-Acetylgalaktosamin-Serine/Threonine, welches die Kernstruktur der Mucin-Protein *O*-Glykosylierung darstellt. Das T-Antigen kommt in vielen Tierarten vor. Während es in Säugetieren wahrscheinlich nur eine T-transferase gibt, zeigten in *Drosophila melanogaster* vier Isoenzyme die entsprechende Aktivität. Transkripte des aktivsten dieser Enzyme in *Drosophila melanogaster* wurden in der maternalen mRNA und in der Amnioserosa gefunden. Die Interaktion zwischen Amnioserosa und Epidermien wird als Wundheilungsmodell betrachtet, trennt sie doch während der Embryonalentwicklung zwei Epidermien, die laterale und die dorsale, die nach dem Abbau der Amnioserosa zusammenwachsen. Die mRNA zweier anderer Enzyme wurde in den Speicheldrüsen des Embryos gefunden. Die Rolle dieser Enzyme in der Embryonalentwicklung und Physiologie von *Drosophila melanogaster* muss erst noch mit Hilfe von *in vivo* Modellen untersucht werden. Mit dieser Arbeit liegen die enzymatischen Grundlagen dazu jetzt vor.

GLYCANS AND GLYCOSYLTRANSFERASES OF DROSOPHILA MELANOGASTER

1 Summary

Glycans are the carbohydrate-moiety of glycoproteins and glycolipids. Their function ranges from the modulation of specific ligand-receptor interactions to providing stabilizing and barrier functions. Of all genes responsible for the formation of glycan chains, glycosyltransferases are the foremost targets for the genetic analysis of glycan function since they are usually specific for one glycan-type and the linkage produced. One structurally related family of glycosyltransferases catalyzes the elongation of glycans with monosaccharides in a β1,3 linkage. In *Drosophila melanogaster*, β1,3 glycosyltransferases have been shown to be implicated in the synthesis of glycosphingolipids, proteoglycans, mucin type O-glycans and O-fucose initiated O-glycosylation of proteins carrying epidermal growth factor receptor (EGFR)-modules. This dissertation discusses some principles of biosynthesis, structure and function of glycans in *Drosophila melanogaster*.

2 Definitions

Glycans are the carbohydrate moiety of glycoproteins or glycosphingolipids, which belong to the super-family of glycoconjugates. Glycoconjugates that have lost their glycan chains are called aglycones. Glycans contain one to many monosaccharides linked together as oligomers or polymers (Figure 1).

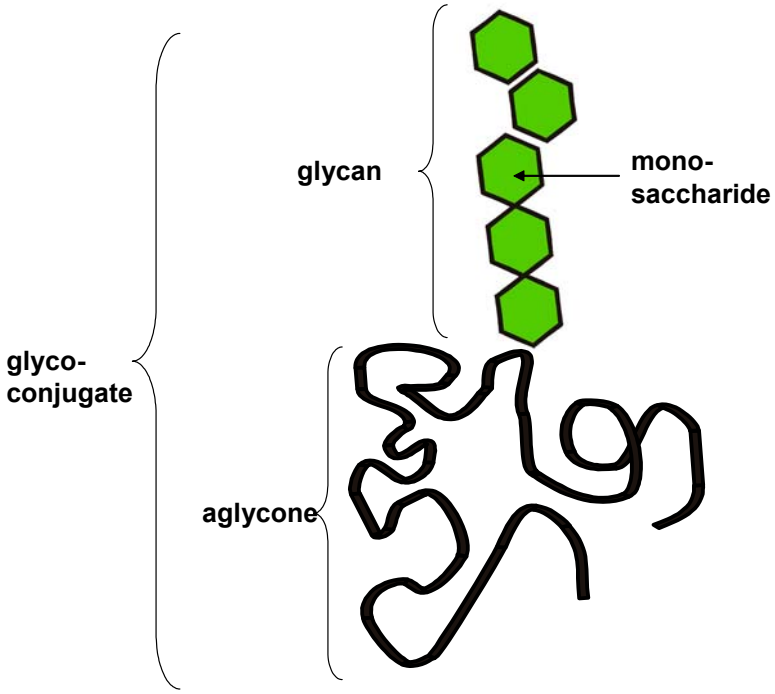

Figure 1: Definitions for glycoconjugate, glycan, aglycone, and monosaccharide.

A glycan is termed "complex" if it contains more than one type of monosaccharides. Complexity is no structural feature since a simple glycan like cellulose, consisting solely of glucose, may yield highly complex three dimensional structures.
The common monosaccharide residues in animal glycans are glucose (Glc), galactose (Gal), mannose (Man), N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), fucose (Fuc), xylose (Xyl), glucuronic acid (GlcA), iduronic acid (IdoA) and sialic acid (Sia). Examples of linear representations of their chemical formula are shown in Figure 2. Monosaccharides in glycans of animals are either pentoses or hexoses and their derivatives. The carbons are numbered according to the rules of organic chemistry with C-1 to C-6 (Figure 2). Two

monosaccharides are referred to as “epimeric”, if the stereochemistry at a single chiral C-Atom differs. For example, Man is the C-2 epimer of Glc, whereas Gal is its C-4 epimer (Figure 2). Monosaccharides may derive from N-acetylation at C-2, yielding for example the hexosamines GlcNAc and GalNAc, or from oxidation at C-6, producing the uronic acids GlcA and IdoA. Dehydroxylation at C-6 yields the deoxyhexose Fuc. For an overview of the many types of sialic acids found in animal glycans the reader is referred to (1). The orientation of the hydroxyl group at C-5 in hexoses (C-4 in pentoses) defines the absolute D- or L- configuration of the monosaccharide. D-monosaccharides predominate in nature. Monosaccharide residues within glycans may also contain non-carbohydrate moieties, as for example sulfate groups (-SO_4), phosphoethanolamine (pEtn) and phosphocholine (pCho) covalently linked via an ester bond to a free hydroxyl group within monosaccharides.

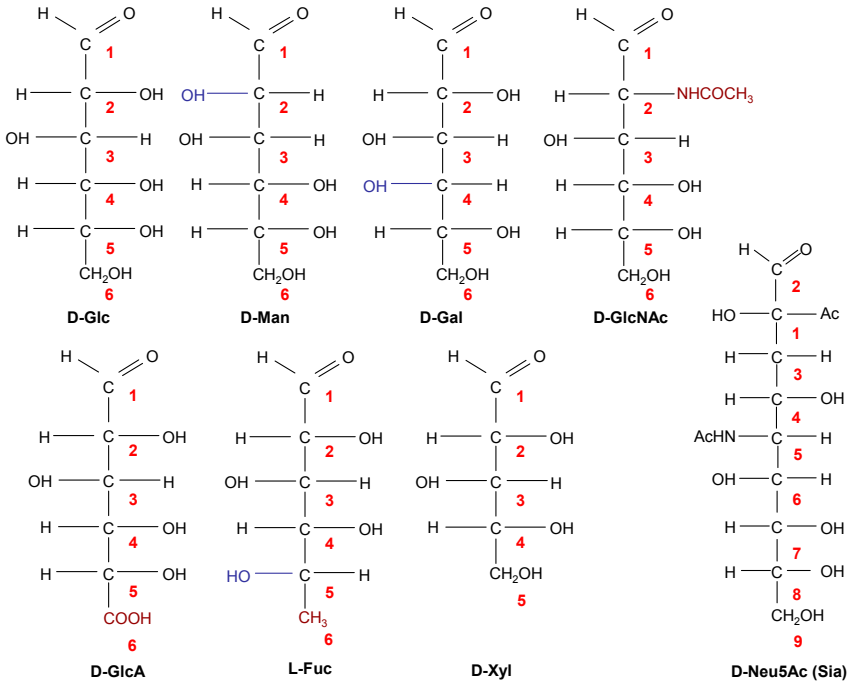

Figure 2: Structures of common monosaccharides in animals. Epimers: Glc and Man are C-2 epimers to each other while Glc and Gal are C-4 epimers. Modifications: GlcNAc is a derivative of Glc which is N-acetylated at C-2, GlcA is a derivative of Glc with its –OH group at C-6 oxidized. Fuc is the only L-monosaccharide commonly found in animal glycans and is a 6-deoxyhexose. Xyl is the only pentose commonly found in animal glycans. There are more than 30 different derivatives of Sia but Neu5Ac is the most common.

Based on the type of functional group besides the hydroxyl-groups, monosaccharides are classified into aldoses (all monosaccharides mentioned above) or ketoses (e.g. fructose). Aldoses contain an aldehyde group (–CH=O) whereas ketoses contain a ketone group (-CR=O). Important chemical properties of the monosaccharide are mediated by the aldehyde or ketone

group. Fore exemple, monosaccharides in solution are not often present as open, but more as cyclic forms. This is the consequence of the formation of an intramolecular bond within the monosaccharide between its ketone/aldehyde-group and one of its hydroxyl groups, creating a four-substituted carbon and thus a new chiral center on the carbon upon cycle formation. This type of reaction is chemically referred to as hemiacetal-formation. The new chiral centre defines the anomericity of the monosaccharide and is traditionally symbolized with α or β. The intramolecular hemiacetal is usually formed with the hydroxyl group at C-5 in hexoses or C-4 in pentoses. The anomeric carbon is also engaged in the covalent bond between monosaccharides. The covalent bond between two monosaccharides is called the glycosidic linkage or glycosidic bond and refers to the oligo- and polysaccharide as the peptide bond does to oligo- and polypeptides, or the phosphodiester bond to oligo- and polynucleotides. To form the glycosidic bond, the α or β anomeric carbon of the hemiacetal monosaccharide engages in a bond with a hydroxyl group of another monosaccharide to form an acetal, protecting the reactive aldehyde or ketone group. Thus, the monosaccharide involved in the glycosidic linkages via its anomeric carbon loses its reduction potential, while the other monosaccharide engaged in the bond with its hydroxyl-group is still in equilibrium with its open form and retains its reducing aldehyde group. The glycosidic linkage of a disaccharide is described by the anomericity (α or β) of the monosaccharide at the non-reducing end and the number of the carbon to which the reacting hydroxyl group of the second monosaccharide is attached. For example, the notation "Galβ1,3GalNAc" means, that the β anomer of Gal at the non-reducing end is engaged in the glycosidic bond with its anomeric carbon while GalNAc is engaged in it with the hydroxyl group at C-3.

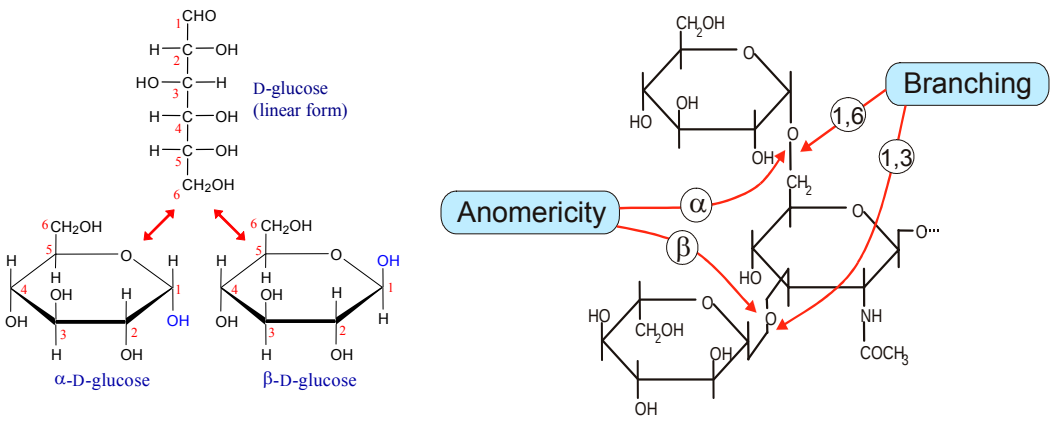

Figure 3: **Monosaccharides form rings and glycosidic linkages**. **Left:** Carbons of glucose are numbered from C-1 to C-6 starting at the carbon of the aldehyde group. Glucose forms an intra-

molecular hemiacetal, as the C-1 aldehyde group and the C-5 hydroxyl group react to form a six-membered ring. Cyclization yields two stereochemically different products with respect to C-1, the anomeric carbon. The two anomers of glucose are shown in Haworth projection. **Right:** Glycosidic linkages between two monosaccharides are formed by the hydroxyl group at the anomeric carbon (C-1) and any free hydroxyl group in the other monosaccharide. Exemplified are a Glcα1,6GlcNAcα and a Glcβ1,3GlcNAcα linkage.

While the peptide bond only yields two non-equal possibilities to link two amino acids, i.e. either Ile-Leu or Leu-Ile, two different hexoses may yield 16 different disaccharides. Chemically, more than $1.5x10^{12}$ structures could be generated from six hexoses (2). However, only very few of these linkages are formed *in vivo*, because glycans are made by enzymatic synthesis and one enzyme usually only catalyzes one linkage. For example, Gal is only found in α or β1,4 and α or β1,3 linkages in animals (while bacteria and fungi also display α and β1,6 linkages) (3). These linkages are formed by glycosyltransferases (GTs), or - in a few cases - by glycosylhydrolases that perform trans-glycosylation (4). The enzymatic formation of different glycoconjugates is called glycosylation.

3 Glycoconjugate classes

The process of glycosylation yields different classes of glycoconjugates (Figure 4). In animals, glycans can be found attached to ceramide, yielding glycosphingolipids or they are protein-bound, yielding glycoproteins. The different types of glycans can be recognized by their core structures, which are conserved among most animals with small variations, while elongated structures may vary between species, cell type, tissue, developmental stage or other determinants.

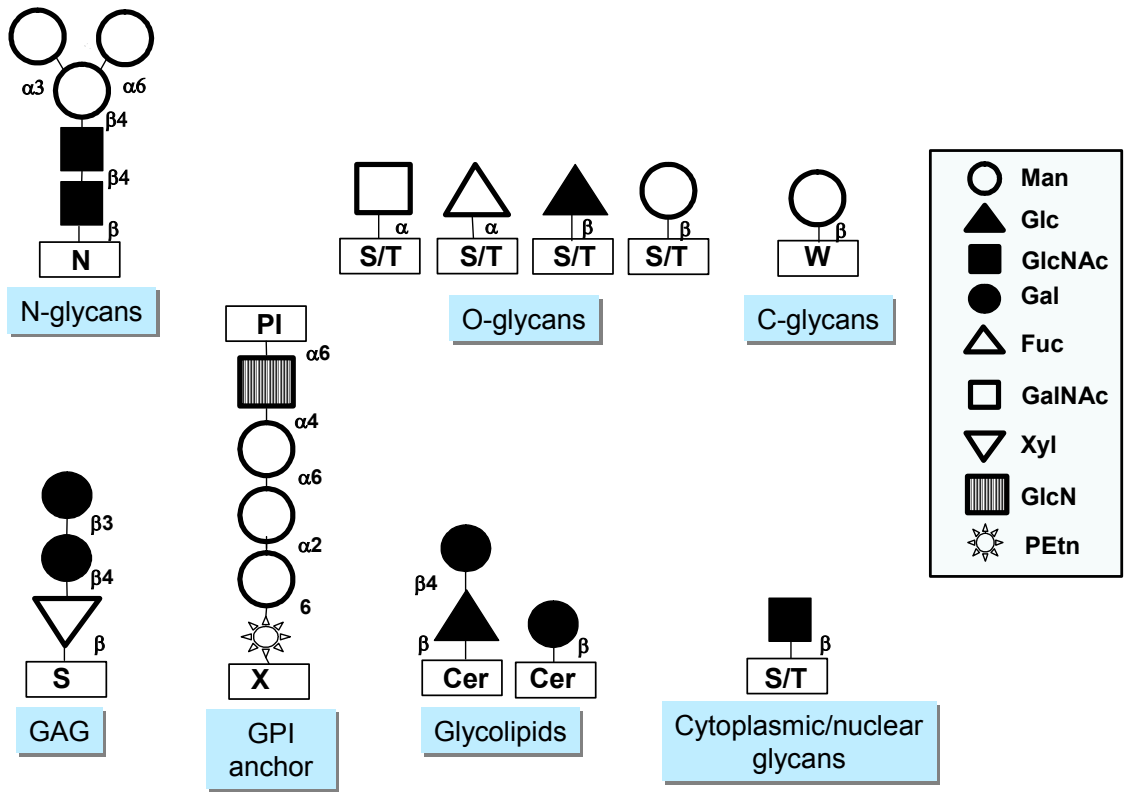

Figure 4: Core structures of eukaryotic glycans. ER N-glycans are preassembled and attached *en bloc* to the amine group of asparagines. Monosaccharides bound via the -OH groups of serines or threonines constitute the cores of the different O-glycans, αGalNAc attached O-glycans are also called "mucin-type". In C-glycans, Man is bound to tryptophan via C-2 of the indole ring, the glycosaminoglycan (GAG) - core structure is attached to -OH groups of serines but not threonines. GPI anchor- glycans are preassembled on phosphoinositol (PI) and then transferred *en bloc* involving a transamidation reaction that results in the cleavage of a signal peptide. The glycan moiety of eukaryotic glycosphingolipids is attached to ceramide (Cer). A list with glycan types encompassing bacterial and plant glycans can be found in (5).

4 Glycan functions

It is difficult to predict the functions mediated by specific glycans or their relative importance to the organism *a priori* since their biological roles "span the spectrum from trivial to crucial for development and survival of an organism" (6). Functions of glycans include the mediation of structural integrity of glycoconjugates or the modulation of specific recognition events (Table 1). The glycan moiety of glycoconjugates is recognized by inter- or intracellular binding partners. Proteins, that bind to carbohydrates are called lectins. Lectins can be proteins of many functional classes, including receptors, signalling proteins, enzymes and structure-preserving proteins. Glycans influence the function of a glycoconjugate in ways that are organism-, celltype-, glycan-type- and glycoconjugate- specific. This may be exemplified by the interaction of the lectins calnexin and calreticulin with N-glycans in the ER for the quality control of glycoprotein-folding (7). Inhibitors of N-glycan biosynthesis have been successfully used to block the exit of N-glycosylated hepatitis B virus proteins from the ER to the secretory pathway (8), leading to virus aggregates in the ER and substantial decrease of the organism's viral load. The inhibition of as little as 6% of cellular glycoprocessing results in a greater than 99% reduction in the secretion of hepatitis B virus (9), indicating a higher sensitivity of virus proteins to N-glycan synthesis inhibition.

The importance of a particular glycan may also vary between different animal species. For example, absence of the glycosphingolipid GlcNAcβ1,3Manβ1,4Glcβ-ceramide is lethal for *Drosophila* during development while *C. elegans* survives without this glycosphingolipid [I, II]. A main motor in the generation of glycan diversity and function between evolutionary lineages seem to be parasite – host interactions. A plethora of bacteria and viruses bind to glycans (Table 1). The parasite genome evolves towards more specific and stronger binding of host factors and escaping its immune reactions, while the host evolves towards masking or deleting the carbohydrate binding sites used by the invader. Therefore, some host-glycans may have lost their ancient physiological binding partners and functions over time.

Table 1: Diversity of glycan functions

functional category	Examples
protective, stabilizing, organizational, and barrier functions.	The glycocalyx represents the outer protective layer of the cell. Glycans attached to matrix molecules and proteoglycans mediate interactions that are important for the maintenance of tissue structure, porosity, and integrity. Densely O-glycosylated proteins such as mucins produce a physical barrier on surfaces. Solubility of excreted proteins may also be increased by addition of glycans.
susceptibility to proteases	Glycans protect proteins from proteolytic degradation. To overcome O-αGalNac barriers (mucin type O-glycans) in *Trichoplasia ni*, baculoviridae produce a protease that specifically degrades O-glycans (10).
protein folding in the ER	The ER chaperones calnexin and calreticulin bind to incorrectly folded proteins in the ER through mono-glucosylated N-glycans and retain the glycoprotein in the ER, where it is refolded (reviewed in ref. 7).
protein sorting in polarized cells	O-glycans (O-αGalNac) have been shown to be a necessary factor for targeting specific proteins to apical membranes of polarized cells (reviewed in ref. 11).
trafficking and targeting to lysosomes	Man-6-phosphate on N-glycans is a signal that targets proteins to lysosomes (reviewed in ref. 12). Exogenous enzyme can be targeted to lysosomes of macrophages from enzyme deficiency patients by terminal Man on N-glycans in patients with Gaucher's disease (reviewed in ref. 13).
recognition of glycans by antibodies	Blood groups A, B and 0 differ in the structure of glycans displayed on their glycoconjugates. These glycans define the major transfusion incompatibilities between humans. 1% of human IgG antibodies are directed against the Galα1,3Gal epitope (14), which is found on glycoconjugates from most animal cells but old world monkeys and primates. These constitute a major xenotransplantation barrier.

recognition of host glycans by invading microbes	N- and O-glycans, glycosphingolipids and proteoglycans can serve as (co-) receptors for many bacteria, viruses and other parasites as *Bordetella pertussis, Giardia lamblia* or *Entamoeba histolytica* to name a few (see as well http://sugarbinddb.mitre.org/)
modulation of binding affinities of receptor and ligand	O-Fuc initiated O-glycans modulate the affinity of the Notch receptor with its ligands Delta and Serrate in *D. melanogaster* (15).
direct receptor response modulation	Tyrosine auto-phosphorylation of the EGF receptor is regulated by the presence of gangliosides, a class of glycosphingolipids (reviewed in ref. 16)
range or specificity of hormone function	Glycosylation variants of human thyroid-stimulating hormone have physiological properties that have been shown to selectively activate signal transduction through either the cAMP or inositol phospholipid hydrolysis pathways (17).
range or specificity of growth factors	Heparan sulfate proteoglycans shape gradients of growth factors/morphogens like the FGF-, TGFβ- and the WNT family.
intracellular signalling	Several intracellular signalling and chromatin proteins carry O-β–GlcNAc, including RNA polymerase II transcription factors. Functions include nuclear transport, assembly into multimeric complexes, sensoring of cellular energy and regulation of phosphorylation (reviewed in ref. 18).
Secretion rate and circulatory half life	Desialylated N-glycans on erythropoietin (EPO) are trapped in the liver and excreted more quickly, tetra-antennary N-glycans are more active *in vivo* than bi-antennary N-glycans (19). Similarly, sulfate-bearing O-Man initiated O-glycans increase the circulatory half-lifes of pituitary glycoprotein hormones (20).

5 Concepts of glycan synthesis

DNA, RNA and proteins are the products of specific templates, thus a given template DNA sequence will code for a unique amino acid sequence. In contrast, glycans are produced in an "ordered melting pot"or on an "assembly line" (Figure 5). Moderately simplified, a model of glycan formation can be considered to comprise four steps:

1. Synthesis of activated monosaccharides in the cytosol from precursor-substrates by a large biosynthetic machinery (see section 6.1).
2. Transport of activated monosaccharides by specific transporters into the subcellular compartments where glycosylation reactions take place (see section 6.2).
3. Catalysis of the first glycosylation reaction by initiating GTs. Thereby a mono- or an oligosaccharide is transferred onto proteins or lipids (see section 6.3.1).
4. Elongation of the glycan (see section 6.3.2). Thereby, elongating GTs utilize the product of the preceeding GT as acceptor-substrate.

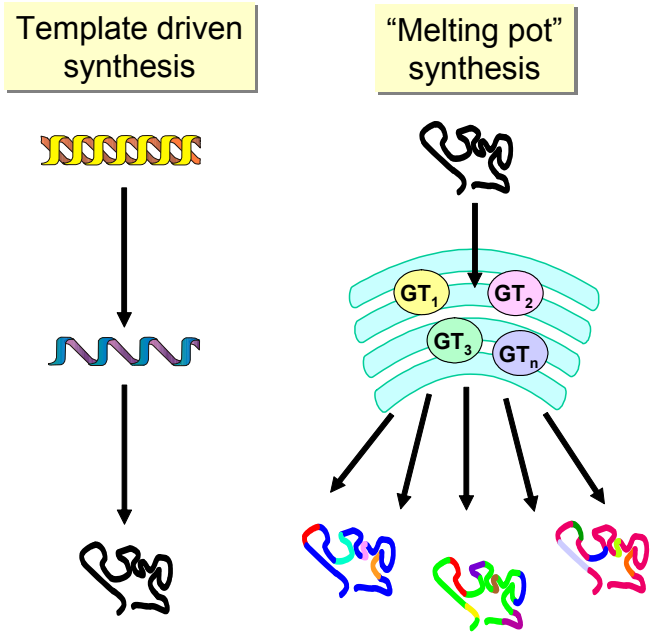

Figure 5: Outline of glycan synthesis. Left: Template driven synthesis of proteins. DNA is the linear template for the corresponding RNA, which in turn is the template for the synthesis of a protein. **Right:** The glycan produced by a cell is determined by the presence of the corresponding GTs, transporters and the type of activated monosaccharides. The glycan may be elongated in different ways, depending on which GT is present in the cell and on the presence of the acceptor and donor-substrate.

Unlike DNA, RNA and proteins, glycans can also be branched. A proofreading system for glycans does not seem to exist. Differences in the cellular set of GTs produce microheterogeneity of glycans within the same glycoconjugate as well as differences between tissues, developmental stages, neoplastic conditions and mutants for parts of the glycosylation machinery. The composition of glycans also depends on the exact localization of GTs along the secretory pathway. Taken together, the synthesis of glycans and the consistency of the glycan structures produced in a cell rely on the presence of the activated monosaccharide and on the presence of the GT at its correct location.

6 The glycosylation machinery

6.1 Activated monosaccharides

Carbohydrates are taken up as nutrients by transporters or synthesized within the cell. In eukaryotes, there are two types of transporters that translocate their substrates from outside the cell into the cytoplasm. One type includes energy-independent facilitated diffusion transporters, such as the family of GLUT proteins (*glu*cose *t*ransporter) (21). A second type of carbohydrate transporters is energy-dependent. The family of *s*odium-*glu*cose co-*t*ransporters (SGLT) (22) is found for example in epithelial cells of the intestine where absorption of nutritional carbohydrates takes place. Cellular glycans constitute another source of monosaccharides. Salvage pathways are fuelled by the disintegration of cellular glycans, which occurs at low pH in the lysosomes by glycan specific endo- and exoglycosidases, thereby liberating monosaccharides from glycoconjugates. Absence or inactivation of glycosidases may lead to a severe imbalance in the carbohydrate metabolism and lead to storage disorders (reviewed in ref. 23).

Table 2: activated monosaccharides

monosaccharide	activated form
Glc	UDP-Glc, DolP-Glc
Gal	UDP-Gal
GlcNAc	UDP-GlcNAc
GalNAc	UDP-GalNAc
GlcA	UDP-GlcA
IdoA	UDP-GlcA[1]
Xyl	UDP-Xyl
Man	GDP-Man, DolP-Man
Fuc	GDP-Fuc
Sia	CMP-Sia

[1]IdoA is epimerized from GlcA after transfer to the glycan chain (c.f. section 8).

Activated monosaccharides are produced in the cytoplasm, with the exception of CMP-Sia, which is synthesized in the nucleus and exported to the cytoplasm. Once the sugars are available as monosaccharides, they are activated by the addition of high-energy groups, these are mono-/diphosphonucleotides (UDP, CMP and GDP) and dolichylphosphate. Dolichylphosphomannose (DolP-Man) and dolichylphosphoglucose (DolP-Glc) are synthesized from the nucleotide

activated sugars GDP-Man and UDP-Glc and are used in the ER for the synthesis of O-mannosylated, C-mannosylated and N-glycosylated glycoconjugates, as well as for steps in the GPI-anchor synthesis. The activated monosaccharides used for glycan synthesis by animals are summarized in Table 2. The synthesis of activated monosaccharides is regulated by feedback-inhibition or activation of enzymes. These regulatory steps prevent the organism from consequences of an imbalanced uptake of monosaccharides (Figure 6).

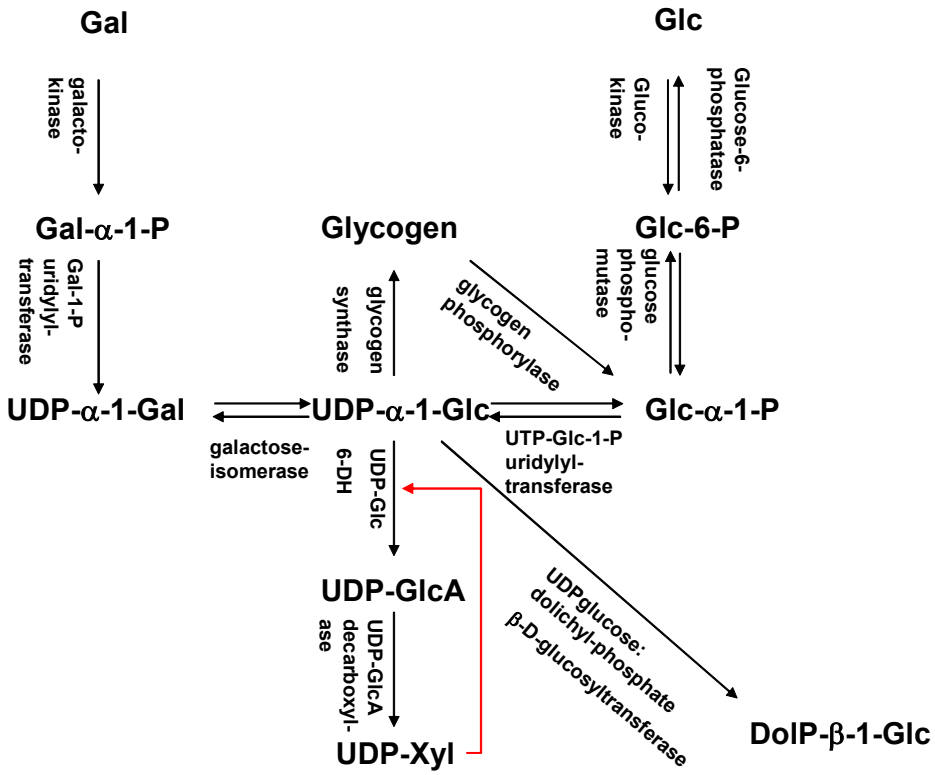

Figure 6: **Excerpt of the activated monosaccharide synthesis pathway**. Glucose is the central metabolite for *de novo* synthesis of activated monosaccharides. All activated monosaccharides can be synthesized from glucose. Storage and synthesis of glycogen, the cellular storage form of glucose, is additionally regulated according to the energy requirements of the cell. Synthesis of an activated monosaccharide may be controlled by feed-back inhibition as exemplified by UDP-Xyl inhibiting UDP-Glc-6-dehydrogenase (red arrow). Figure modified from (24).

6.2 Intracellular transport of activated monosaccharides

Glycosylation reactions primarily take place in the lumen of the ER and the Golgi apparatus, while the synthesis of activated monosaccharides is located in the cytosol. Therefore, activated monosaccharides need to be translocated into these organelles. The transport of nucleotide activated monosaccharides is carried out by specialized proteins, nucleotide sugar transporters. Several of these transporters have been cloned and are located in the membranes of ER and/or

Golgi, depending on whether their transported substrate is required there (25, 26). Maintenance of the concentration of activated monosaccharides in the organelle is also achieved by transporters, since they have been found to act as antiporters, exchanging monophosphorylated nucleotides (products of glycosylation reactions and a subsequent nucleotide diphosphatase reaction) with nucleotide activated monosaccharides in equimolar ratio (27). For example, several transporters of UDP-nucleotide sugars have UMP as a common antiporter (28).

6.3 Glycosyltransferases

Once the activated monosaccharide has reached the site of glycan biosynthesis, they are utilized by glycosyltransferases (GTs) to glycosylate proteins or lipids. GTs generate a glycosidic bond between the anomeric carbon of an activated saccharide and an acceptor. GTs can be classified into initiating (glycosylating aglycones) and elongating (glycosylating glycoconjugates on their glycans) enzymes.

Donor substrates for glycosylation-reactions are mono- or diphosphonucleotide-α–monosaccharides, DolP-β-monosaccharides or a DolPP-activated oligosaccharide (see section 6.3.1 and 8.1). If the transferred monosaccharide retains the anomeric configuration of the donor monosaccharide (β for DolP – monosaccharides and α for nucleotide-monosaccharides), the transferase mechanism is defined as retaining, while GTs that change the anomericity of the transferred monosaccharide are inverting GTs (29). GT reactions are specific as they normally catalyze the formation of one linkage between a specific donor and a specific acceptor only (glycan, protein domain or ceramide), although there are some exceptions to this rule (30, 31). The “one enzyme - one linkage” concept (32) implies the presence of a large number of GT encoding genes, given the diversity of glycans. Indeed, it is estimated that over 200 GTs exist in the mammalian genomes (33) and that some 0.5 to 1% of the human genome code for constituents of the glycosylation machinery (34).

As a first approximation, post-translationally acting GTs are suspected to be arranged in an assembly line in the Golgi, where early acting enzymes are more concentrated in the *cis*-Golgi, intermediate acting enzymes in the *medial*-Golgi, and those adding terminal structures in the *trans*-Golgi. They may, however, also be found on the cell membrane or secreted into body fluids (35) for unknown reasons, where the concentration of activated monosaccharides is too low to carry out transferase reactions. The catalytic domain is sometimes separated from the N-terminus of the enzyme by cleavage through proteases near the stem region (36, 37), which might be another step in the regulation of glycan production.

Failure to localize a GT correctly may result in the loss of the glycan chain (38). Indeed, a mutation in the COG7 gene encoding a Golgi vesicle protein has been shown to result in Congenital Disorder of Glycosylation -IIe (39). Other mechanisms besides those involving trafficking factors may also apply to localize GTs. A first model predicts that the length of its transmembrane domain determines the degree of *trans*-Golgi localization, since membrane thickness increases with a gradually higher cholesterol and glycolipid concentration from *cis*- to *trans*- Golgi (40). Another hypothesis is that inclusion of GTs into constitutively formed secretory vesicles is blocked upon di- or oligomerization of GTs and that GTs therefore remain in the Golgi (41). Indeed, Golgi localized GTs have been shown to form dimers and other complexes *in vivo* and *in vitro* (42-47).

6.3.1 Initiating Glycosyltransferases

Initiating GTs add the first unit of a glycan chain to nascent or folded proteins in the ER or as post-translational modifications in the Golgi. Their acceptor substrates are proteins, ceramide or dolichophosphate and their donors may be nucleotide activated monosaccharides, lipid activated monosaccharides or the preassembled lipid-linked oligosaccharide $Glc_3Man_9GlcNAc_2$ (Table 3). For protein glycosylation, the sequence context of the amino acid to be glycosylated is important as well as the 3D structure and the accessability of the site within the protein. Thus, not all amino acids within potential glycosylation motifs are glycosylated. On the other hand, some O-α-GalNAc glycosylation initiating GTs require peptides, which are glycosylated at an adjacent site (48). Glycosphingolipids are formed by attachment of Glc or Gal to ceramide at the cytosolic face of the Golgi Apparatus (49, 50), implying, that the product is flipped into the lumen of the Golgi before elongation can occur. Multi-enzyme complexes which transfer preassembled glycans onto proteins are present in the N-glycosylation pathway and in the GPI anchor synthesis. In the N-linked pathway, the glycan is preassembled on a lipid carrier and the dolichylpyrophosphate-linked oligosaccharide $GlcNAc_2$ Man_9 Glc_3 is transferred *en bloc* onto proteins by a multisubunit protein complex in the ER membrane to asparagine residues on nascent proteins. This complex is termed the oligosaccharyltransferase (c.f. section 8.1). Differently, preassembled GPI anchors from the cytoplasmic side of the ER (pEtn-Man_3-GIcN-acyl-P-lnositol) are flipped into the ER lumen und transferred by a large enzyme complex *en bloc* onto proteins by a transamidase-reaction.

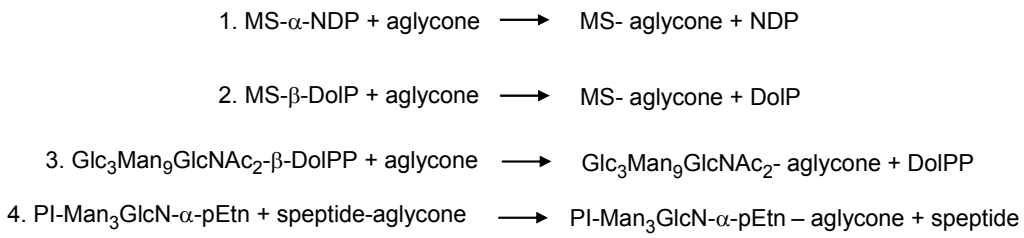

Figure 7: Initiating reactions. Activated monosaccharides (MS) or preassembled activated oligosaccharides are transferred onto proteins or lipids by initiating GTs (reactions 1-3). GPI-anchor modification of proteins (reaction 4) is not a GT reaction since it involves transamidation and hence cleavage of a signal peptide (speptide).

Table 3: Glycan types and their initiating glycosyltransferase reactions in animals

Glycan type	acceptor, peptide sequence[1]	Donor substrate	Localization of transfer
Lipid linked oligosaccharide	DolP	UDP-GlcNAc (to give DolPP-GlcNAc)	cytosolic face of ER
N-glycans	Asn-X-Ser/Thr (X ≠ Pro)	$Glc_3Man_9GlcNAc_2$-DolPP	ER-lumen
O-glycan types:			
O-GalNAc	Ser/Thr [2]http://www.cbs.dtu.dk/services/NetOGlyc/	UDP-GalNAc	cis-Golgi
O-Glc	$Cys_1XxxSerXxxProCys_2$ (51)	UDP-Glc	?
O-Fuc		GDP-Fuc	ER (53)
EGFR	$Cys_2XxxXxxGlyGlySer/ThrCys$ (51)		
TSP	$TrpXxx_5CysXxx_{2/3}Ser/ThrCys Xxx_2Gly$ (52)		
O-Man	?	DolP-Man	ER (54)
Proteoglycans	variations of SerGly motifs (55)	UDP-Xyl	ER and/or cis-Golgi (56)
Glycosphingo-lipids	Ceramide	UDP-Glc, UDP-Gal	cytosolic face of Golgi (49)
C- Man	TrpXxxXxxTrp	DolP-Man	ER? (57)
O-GlcNAc	[2]Ser, Thr http://www.cbs.dtu.dk/services/YinOYang/	UDP-GlcNAc	nucleus, cytoplasma
GPI anchors	[2]http://mendel.imp.univie.ac.at/SEQUENCES/gpi-biosynthesis/ and (58)	GPI lipid anchor[3]	cytosolic face of ER

[1]sequences are still being refined, [2]complex prediction and computational methods as neural network algorithms are preferred,[3]varies strongly between species, often: PI-Man_3GlcN-pEtn.

6.3.2 Elongating glycosyltransferases

Elongating GTs in the Golgi catalyze the transfer of monosaccharides from mono- (Sia) or diphosphonucleotides (all other monosaccharides) onto glycans consisting of one or several monosaccharides. Elongating GTs at the ER membrane are primarily involved in the assembly of the N-glycan precursor lipid and the GPI-anchor.

The linkages produced classify elongating GTs. For example, β1,4 GTs form a linkage between the β anomeric carbon of the transferred monosaccharide and the C-4 attached hydroxyl group of the acceptor monosaccharide. There are α and β 1,2, 1,3, 1,4, 1,6 hexosyl- and pentosyl-transferases (Table 4). The anomeric carbon of Sia is C-2 and linkages produced by SiaTs are described accordingly.

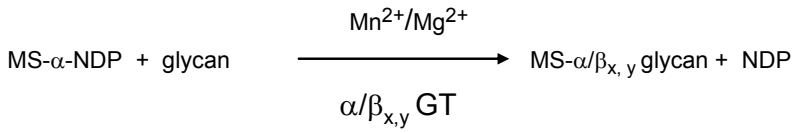

Figure 8: **Elongating GTs** catalyze the transfer of a monosaccharide (MS) from an activated MS onto a glycan in a specific linkage. Transfer occurs, depending on the specificity of the GT, onto α or β anomeric terminal monosaccharides. Most GTs display absolute requirement of Mn^{2+} as co-factor, exceptionally, Mg^{2+} is the necessary co-factor.

Sequence homologies between GTs that catalyze a specific linkage are higher than homologies between GTs using the same donor or acceptor substrates. This means that a β1,3 GalT is more homologous to a β1,3 GlcNAcT than a β1,4 GalT to a β1,3 GalT ((3) and section 9, Figure 21) indicating that the donor substrate is more easily changed than the linkage for most families. This rule does not apply to FucTs and SiaTs (Figure 10), whose sequence similarity is higher among each other than among enzymes transferring the same D-hexoses.

Even if there are many homologous regions between members of a GT-family, sequence conservation at the DNA level is low, which has limited the use of DNA hybridisation techniques to the cloning of sialyltransferases (59). Therefore, computational sequence comparison at the protein sequence level (60) is used for the identification of cDNAs putatively encoding GTs that generate a specific linkage while donor and acceptor specificities of the transferase are less easily determined (47, 61).

Table 4: Elongating glycosyltransferases acting on glycans of animal proteins or glycosphingolipids

donor substrate	linkages catalyzed	[1]EC numbers
UDP-GlcNAc	β1,3 β1,4 β1,6 α1,4 β1,2	**2.4.1.149**, –146, –147, –163 **2.4.1.144**, –145, –197 **2.4.1.102,** –148, –150, –164, –115 **2.4.1.223**, –224 **2.4.1.143**, –101
UDP-GalNAc	β1,4 α1,3 β1,3 β1,6	**2.4.1.92**, –165, –174, –175 **2.4.1.40**, –88 **2.4.1.79** 2.4.1.154[2]
UDP-Glc[3]	α1,4 α1,3	**2.4.1.186** UGGT
UDP-Gal	α1,3 α1,4 β1,4 β1,3	**2.4.1.87**, –37 **2.4.1.228** **2.4.1.38**, –86, –22 **2.4.1.122**, –179, –62, –134, –179
UDP-GlcA	β1,3	**2.4.1.135**, –226
UDP-Xyl	-	[4]n.d.
GDP-Man[5]	α1,2 α1,3 β1,4 α1,6	2.4.1.131 2.4.1.132 2.4.1.142 2.4.1.232
GDP-Fuc	α1,3/1,4 α1,2 α1,3 α1,6	**2.4.1.65** **2.4.1.69** **2.4.1.214**, –152 **2.4.1.68**
CMP-Sia	α2,8 α2,3 α2,6	**2.4.99.8** **2.4.99.4**, –2, –5, –6, –7, –9, –10 **2.4.99.1**, –3, –11

[1]enzyme EC numbers were compiled from: http://www.brenda.uni-koeln.de.
[2]not cloned, linkage based on data from purified enzyme analysis, no sequence data available.
[3]GlcTs ALG 6, 8 (α1,3) and 10 (α1,2) use DOL-P-Glc as donors and have not yet been assigned EC numbers nor has UDP-Glc: glycoprotein GlcT (UGGT).
[4]in contrary to plants and prokaryotes, no elongating activity with UDP-Xyl has been observed in animals yet.
[5]ALG ManTs are located at the ER membrane and elongate the N-glycan DolPP- precursor oligosaccharide.
in bold: sequences of human (mostly) enzymes that were used in Figure 10 for an alignment.

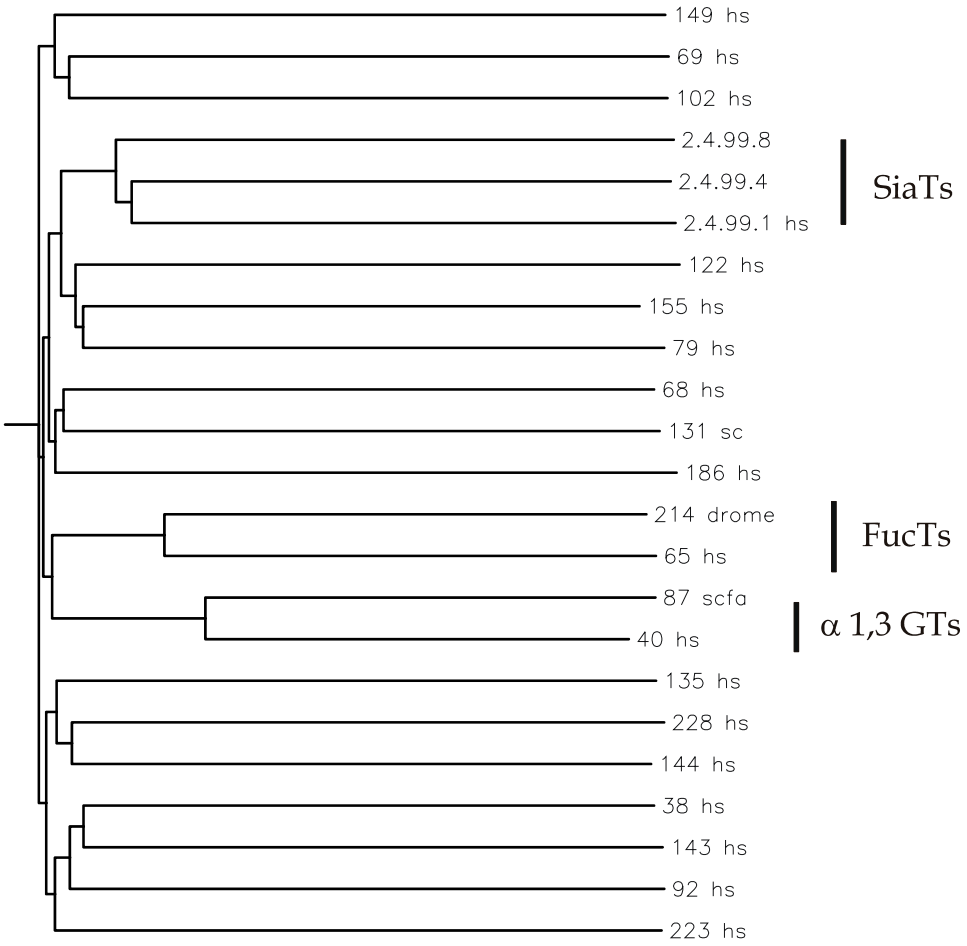

Figure 9: Elongating Golgi GTs catalyzing different reactions. Phylip-tree (62) of a ClustalW-alignment (63)**.** Sequences refer to Table 5. hs: homo sapiens, drome: *Drosophila melanogaster*, scfo: *sus scrofa*.
Alignments show that FucTs and SiaTs, although catalyzing different linkages, display similarities among each other, while at this stage, the only moderately significant sequence similarity between transferases catalyzing the same linkage is found between two α1,3 GTs.

7 *Drosophila melanogaster as a model system to study glycosylation*

Glycosylation has recently attracted more attention because several diseases have been found to be related to glycosylation-defects in humans (64). For the N-glycosylation pathway in the ER, many human-orthologous genes encoding GTs were characterized in yeast mutants (reviewed in ref. 24). However, yeast Golgi glycans differ substantially from those found in multicellular organisms (65). Thus, to identify animal-relevant carbohydrate active enzymes, other procedures like selecting for resistances to cytotoxic plant lectins in mammalian cell lines have been used (66). These biochemical screens yielded a large number of mutant cell lines and led to the identification of transporters of activated-monosaccharides and GTs. In cell lines, phenotypes are more difficult to detect than in multicellular organisms, which often display lethal consequences of disturbed glycan-mediated intercellular interactions. Therefore, model organisms are indispensable to study the functions of specific glycans.

In contrast to other carbohydrate active enzymes, GTs remain the glycobiologist's central target for knock-out models, if elimination of a specific glycan and observation of the systemic effect is desired, because most GTs are specific for both glycan-type and catalyzed linkage. Indeed, reverse genetics approaches with GT-knock-out models in mice have proven powerful tools for the investigation of both physiological and developmental issues in mammals (reviewed in 67). The opposite approach is to look for candidate genes coding for carbohydrate active enzymes by screening the genome of multicellular organisms and phenotyping the mutants - an approach too cumbersome and expensive with mice but proving suitable to find carbohydrate active enzymes in *C. elegans*, for example. A screen for mutations in the vulval development of *C. elegans* revealed genes encoding GTs and transporters that are components of the proteoglycan synthetic pathway (68, 69). Another model system that can be used for screens is *D. melanogaster*, which has the largest number of established mutants of all model organisms. Even more, genetic interactors of GTs may be found by crossing mutants with phenotypically similar ones, or mutants that are otherwise (typically by bioinformatics) presumed to be interactors (70, 71). Analysis of double mutants may reveal phenotypically scoreable interactions in the system under examination.

A genetic interactor could be any glycosylated protein or any that needs to interact with a glycoprotein or glycolipid to carry out its functions. Examples would be specific glycoconjugates, lectins or other carbohydrate active proteins like transporters. Secondary interactors could be

transcription factors, whose signal is integrated in parallel or downstream of a glycoconjugate-mediated function (e.g. 72). Further analysis and epistasis experiments can be employed to determine the level at which the signal is integrated. For example, double mutants for a promiscuous transporter encoded by the gene *fringe connection* with two GTs involved in either proteoglycan or O-glycan synthesis displayed a stronger phenotype than the individual mutants (73), also genetically indicating the cooperation of these carbohydrate active proteins.
Taken together, the wealth of publically available mutants, genetic amenability and a century of experience with this model organism render *D. melanogaster* the organism of choice to search for interactors of glycans.

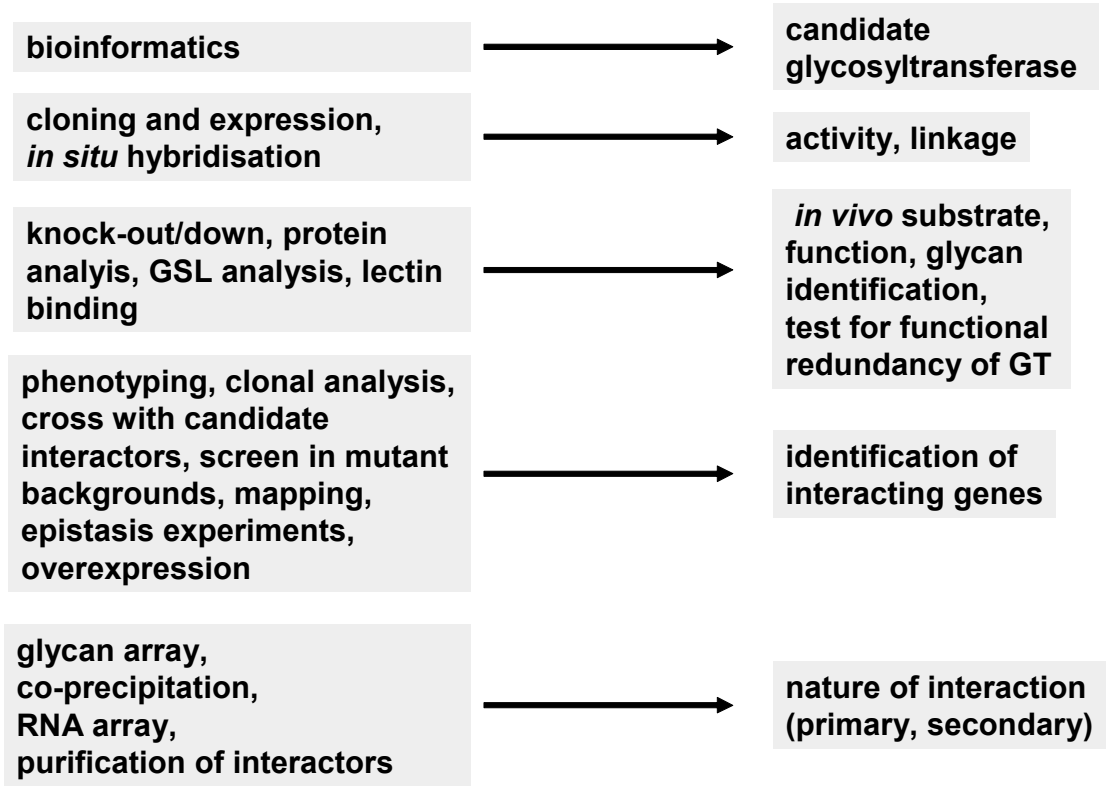

Figure 10: Steps in the analysis of glycan functions. A broad area of techniques encompassing genetic analysis as well as biochemical methods and glycan structure determination is required to find and describe interactors of glycans.

8 Drosophila melanogaster glycans

The first glycosylation mutant described as such in *D. melanogaster* was found in a screen scoring the binding of a CNS specific antibody directed against a glycan epitope. Absence of binding in the mutant CNS indicated a difference in the glycan structure of the mutant. The locus was dubbed *neuronally altered glycosylation* (74). While that gene lacks molecular characterization to date, several glycosylation pathways and glycoconjugate classes have been described in *D. melanogaster* (Table 5). *D. melanogaster* has been shown to carry out N-linked glycosylation, GPI anchor synthesis, O-Fuc initiated O-glycosylation, O-mannosylation, mucin-type O-glycosylation, glycosaminoglycan- and glycosphingolipid- synthesis (reviewed in refs. 75 and 76). Glycans may be produced only during specific developmental phases (77), be restricted to a rare tissue or occur in low abundance for other reasons. Thus, methods involving glycoconjugate-purification from whole animals, glycan release and subsequent glycan structure determination cannot be expected to be comprehensive. Cloning and characterization of GT enzymes can give important clues for the identification of novel glycan structures. In the following sections, biosynthesis and functions of some already characterized glycan classes are discussed.

Table 5: Glycoconjugate classes described in *Drosophila melanogaster*

glycoconjugate class	presence in *D. melanogaster*?	monosaccharides in *D. melanogaster* glycans[1]
N-glycans	yes (78, 79)	GlcNAc, Man, Fuc, Glc
O-glycan-types: O-GalNAc O-Glc O-Fuc O-Man	 yes (48), [III] unknown yes (80) yes[2] (81)	 GalNAc, Gal unknown Fuc, GlcNAc Man
Glycosphingolipids	yes (82)	Glc, Man, GlcNAc, GalNAc, Gal, GlcA
C-glycans	unknown[3](83)	
GPI-anchors	yes (84)	no structural studies to date
Glycosaminoglycans	yes (85)	GalNAc, GlcNAc, GlcA, IdoA, GlcN, sulfated monosaccharides
Cytoplasmic/nuclear glycans	yes (86)	GlcNAc

[1]sialic acids may be also be present on *Drosophila* glycans (discussed in ref. 76)

[2]presence of O-Man glycans is inferred from mutants in a O-ManT-homologous enzyme, not from structural studies.

[3]*D. melanogaster* cell lines do not carry out C-mannosylation.

8.1 N-Glycans

Biosynthesis of N-Glycans (Figure 11) starts at the cytoplasmic side of the ER membrane, where stepwise addition of GlcNAc and Mannose by GTs using diphosphonucleotide monosaccharides leads to the assembly of $Man_5GlcNAc_2$ on the high-energy lipid carrier dolichylpyrophosphate (DolPP). The associated GTs have not been cloned from *D. melanogaster*. Their orthologues are called ALG (for asparagine linked glycosylation). Upon action of a flippase (87), this structure is translocated into the lumen of the ER, where it is further elongated by GTs using DolP activated monosaccharides. Oligosaccharyltransferase (OST) transfers $Glc_3Man_9GlcNAc_2$ from their activator DolPP onto asparagines (Table 3) Glucosidases I and II trim the Glc residues on N-glycans. Then, secreted and membrane glycoproteins have to pass the ER protein quality control to enter the Golgi (reviewed in ref. 7). If the protein is not folded properly, UDP-Glc:glycoprotein GlcT (UGGT) functions as a folding sensor and reglucosylates the N-glycan, recognizing the innermost GlcNAc within exposed hydrophobic regions. Upon recognition of the terminal glucose by the lectin chaperones calnexin and calreticulin, the protein undergoes another folding cycle. If, after release, the N-glycosylated protein is folded correctly, the terminal Glc is cleaved off and the protein exits the ER for the Golgi. A protein, which is not folded correctly for an undetermined number of cycles is targeted for proteosomal degradation in the cytosol in a process called ER associated degradation (88). Once arrived in the Golgi, the processing of N-glycans in *Drosophila* has been shown to include trimming up to complete deglycosylation of N-glycan sites (89) and a number of reglycosylation reactions. Recent studies (79) demonstrated the presence of difucosylated N-glycans and suggested that N-glycans in *D. melanogaster* equal vertebrate N-glycans in terms of diversity, although no sialylated structures have been published to date.

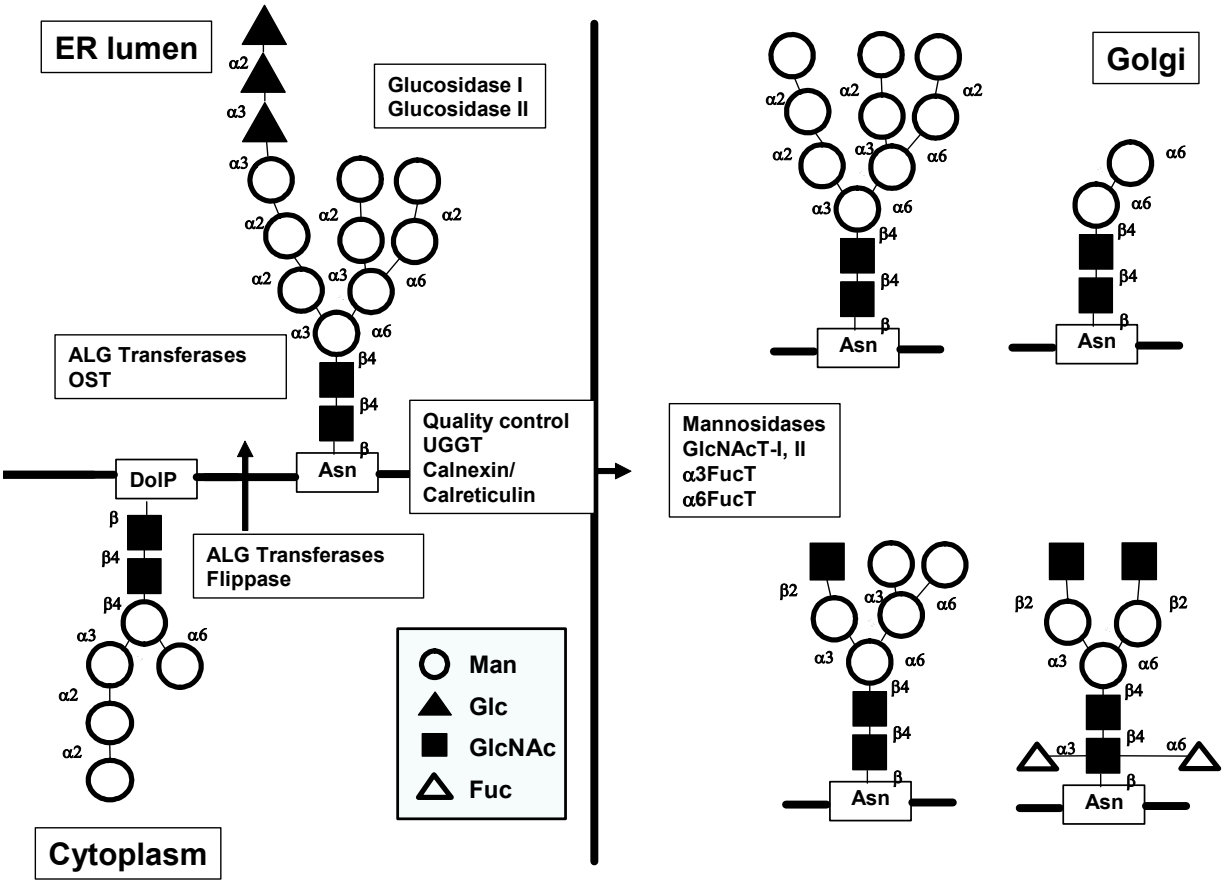

Figure 11: N-glycosylation biosynthetic pathway. N-glycosylation at the ER membrane (left) is conserved in most eukaryotes and involves the action of GTs (ALG, asparagine linked glycosylation) using diphosphonucleotide activated Man and GlcNAc on the cytoplasmatic side of the ER or DolP-activated Man and Glc on the luminal side of the ER to produce a DolPP-assembled glycan. $Glc_3Man_9GlcNAc_2$ is then transferred *en bloc* onto proteins and serves for protein quality control by the lectin-chaperones calnexin and calreticulin. UDP-Glc:glycoprotein GlcT acts as a sensor for incompletely folded proteins. Golgi associated processing (right) is achieved by mannosidases, which trim the high-mannose N-glycan. All structures between $Man_9GlcNAc_2$ to Man_2 $GlcNAc_2$ are present in *D. melanogaster*. GlcNAcT-I and –II are responsible for N-glycan structures with terminal GlcNAc, and at least 2 FucTs for the generation of α3 and α6 core-fucosylated N-glycans.

No systematic analysis on mutations in the N-glycosylation pathway in *D. melanogaster* has been performed to date. Studies on the requirement for N-glycosylation are limited to an embryonic lethal mutation in an ALG3 homologous gene (an α1,3 ManT) encoded by *l(2)not* (90). In humans, a missense mutation in this gene is characterized by microcephaly, severe epilepsy and minimal psychomotor development (91).

8.2 Proteoglycans

Proteoglycans are proteins carrying glycosaminoglycans (GAGs). The GAGs described in *D. melanogaster*, heparan sulfate (HS) and chondroitin sulfate (CS), are linear and unbranched polymers of disaccharides synthesized from an amino sugar and an acidic sugar (Figure 12). They contain regions which are sulfated to various extents. The physico-chemical properties shared with all non-fibrillous polymers of carbohydrates are the capability to bind large amounts of water and to form gels. The negative charge results from the content of acidic sugars and sulfate groups and also allows cation binding. The size of these polysaccharides reaches 100 kD for HS. *C. elegans* vulval development is severely disturbed in mutants for GAG synthetic genes because the vulval extracellular space fails to expand, an effect directly related to reduced water uptake of the tissue failing to produce GAGs.

CS consists essentially of repeated units of the disaccharide $[GalNAc\beta1,4GlcA\beta1,3]_n$ while HS is a strongly modified chain with the backbone disaccharide $[GlcNAc\alpha1,4GlcA\beta1,3]_n$. GAGs commonly found in vertebrates like dermatan sulfate $[GlcA/IdoA\beta1,3GalNAc\beta1,4]_n$ and keratan sulfate $[Gal\beta1,4GlcNAc\beta1,3]_n$ have not been described in *D. melanogaster* to date and hyaluronan $[GlcA\beta1,3GlcNAc\beta1,4]_n$ is absent (92). The main GAG produced during *D. melanogaster* adult life and embryonic development is HS. This GAG is also maternally deposited into the embryo (93). CS, on the other hand, is produced primarily during larval stages (93). CS has a lower content of IdoA than HS and the only described sulfation is carried out by a not yet cloned CS 4-O-sulfotransferase (94).While mutations affecting CS synthesis impact embryogenesis in *C. elegans* (95), no *D. melanogaster* mutant specifically affected in CS formation has been described to date.

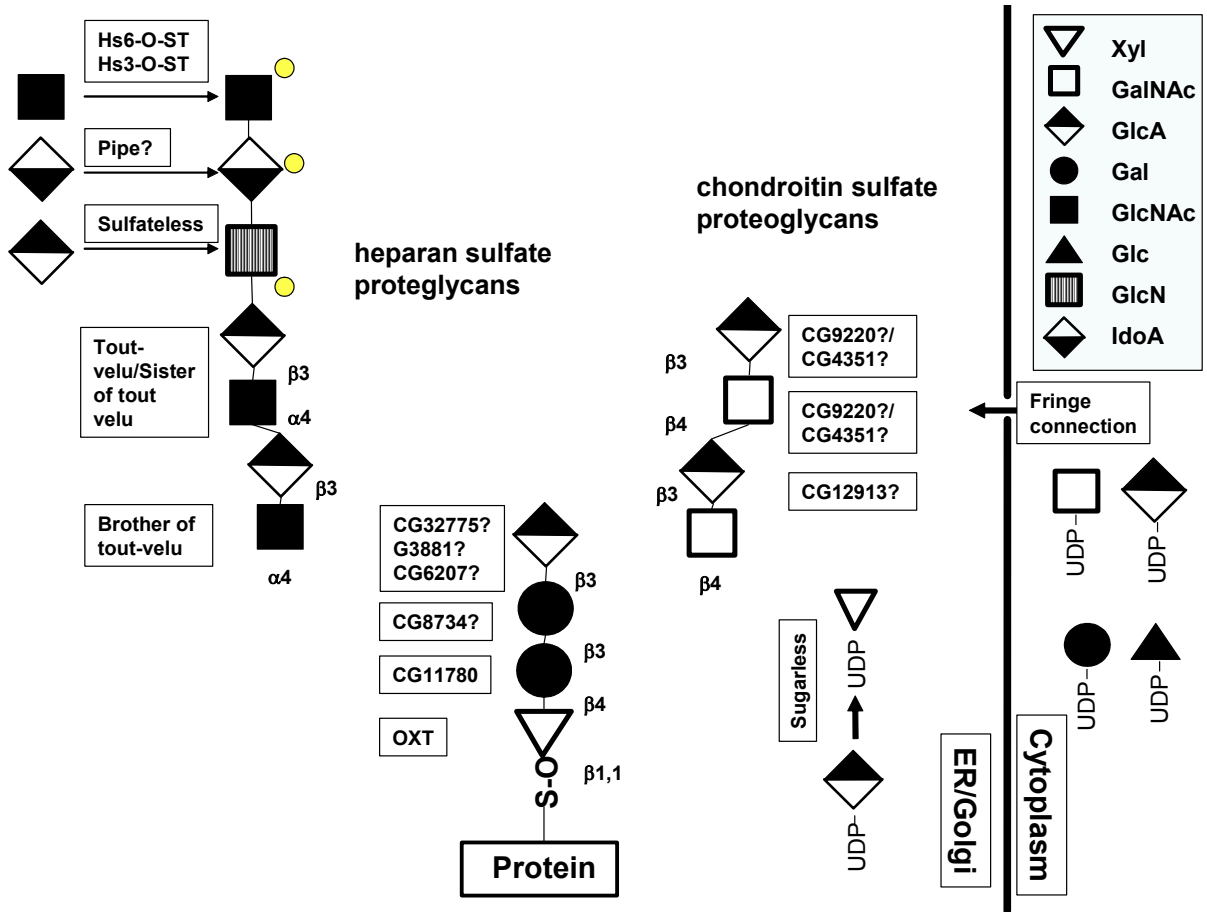

Figure 12: GAG biosynthesis. Enzymes implicated in GAG synthesis or modification are boxed. Question marks after enzyme-names indicate high sequence similarity of the isoform to known human or *C. elegans* enzymes carrying out the same reaction but lacking complete characterization. References for the listed genes are given in the appendix.
Both polymers are assembled in the Golgi from the common linker sequence GlcAβ1,3 Galβ1,3 Galβ1,4 Xylβ1,1 -O-Ser within Ser/Gly- rich regions (96). After the steps leading to the primer tetrasaccharide, an initiating GlcNAcT encoded by *brother of tout-velu* (97) commits the growing chain to the HS fate by a transfer of GlcNAc to GlcA or adds GalNAc to GlcA to synthesize CS. HS synthesis continues by addition of GlcA and GlcNAc. Both activities are catalyzed by a co-polymerase consisting of two enzymes encoded by *sister of tout-velu* and *tout velu* (98). After the synthesis of the backbone structure, the polymer is strongly modified by sulfation and isomerization. These modifications include GlcA epimerization to IdoA, GlcNAc *N*-deacetylation to glucosamine (GlcN) and N-/O-sulfation. Mutants in HS sulfation have been shown to produce strong and/or lethal phenotypes (99-101), substantiating the importance of sulfation within HS chains.

High size, charge and water uptake reduced the perception of GAGs for a long time to a kind of "fly-paper" to which about any ligand would bind with low specificity. However, mutants in GAG biosynthesis paint a subtler picture about the influence of GAGs on the development of *D. melanogaster*, at least as far as HS is concerned. As shown in vertebrates, GAGs may be very specific in their affinity to proteins. In fact, proteins were shown to bind to molecularly defined contiguous sequons in HS (102) and to discontinuous domains forming a loop (103).

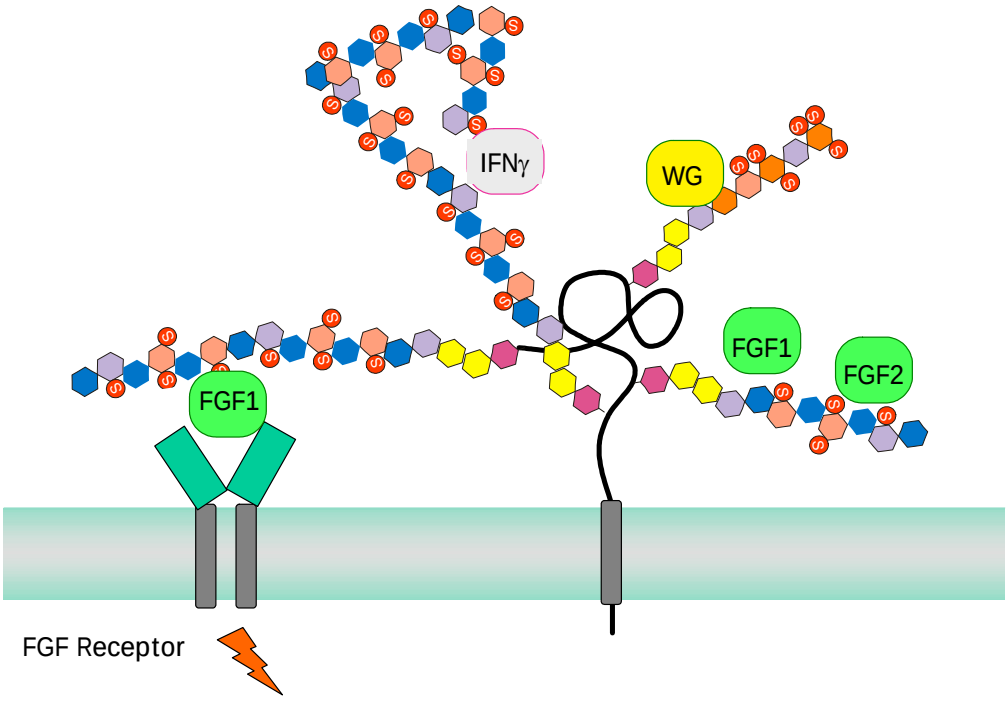

Figure 13: Growth factors and morphogens bind HS as shown in vertebrates and in *D. melanogaster*. Interferon gamma (IFNγ) binds to HS loops and fibroblast growth factors (FGFs) bind different sequons with graded strength. FGF-HS/FGF-1 interaction may fulfil various purposes. For example, modulation of the binding of FGF to HS may regulate FGF-FGF receptor complex formation, receptor dimerization and activation. Growth factor/morphogen-HS interaction may also, in addition to diffusion-control, serve to protect against proteolysis (104). Wingless (WG) also binds to HS and the Wingless gradient is shaped by the glypicans encoded by *dally* and *dally-like* protein. The sequon to which WG binds is unknown, though.

The family of HS carrying proteins, known as heparan sulfate proteoglycans (HSPGs), encompasses glypicans, perlecans and syndecans. Glypicans are HSPGs that are anchored in the plasma membrane by a GPI-anchor. Perlecans and syndecans are HSPGs whose transmembrane domains ensure plasma membrane localization. It seems evident that the glycan moiety is indispensable for HSPG function, since different glypicans, perlecans and syndecans have been shown to similarly influence responses to secreted morphogens like FGF, Decapentraplegic, Wingless and Hedgehog (105-110), although their core proteins differ.

Morphogens are secreted signalling molecules that form gradients in the embryo and in tissues. In short, morphogens are released by a distinct group of cells and are spread into the surrounding area. There, the concentration is differentially interpreted by the receiving cells and induces the expression of a position-specific set of genes (reviewed in ref. 111). For instance, the morphogen Wingless acts at a short range in the embryonic epidermis, regulating the activity of target genes, while as a long-range morphogen in the wing imaginal disc, it forms a stable gradient with a source localized at the boundary between ventral and dorsal compartments (reviewed in ref. 112).

Binding to HS is believed to influence diffusion and stability of the morphogens, helping to establish and maintain the gradient.

The interaction of the morphogen Wingless with the HS biosynthetic machinery has been studied extensively. Direct interaction of HS with Wingless has been demonstrated in cell culture assays (105, 113), and mutant flies in HS chain polymerization display a disturbed Wingless gradient (114). Moreover, they display *wingless* loss-of-function phenotypes (98, 115) as do mutants in activated monosaccharide synthesis (116), in transporters of activated monosaccharides (73, 117) and in HS chain modifying enzymes (99-101, 107).

Since individual morphogens also interact among each other genetically, it can be difficult to assess the relative importance of individual glycan-morphogen interactions. For example, Hedgehog and Wingless functions are interconnected by a feedback-loop during segment polarity determination, a process during which Hedgehog maintains Wingless transcription (118, 119). Morphogen-response modulation by GAGs does not seem to be limited to *D. melanogaster*, since mouse knock-out models in the *tout-velu*-orthologous gene *Ext1* display embryonic lethality (120), while mutations in an EXT gene in humans is associated with multiple hereditary exostoses, benign bony tumours at the growth plate (121).

8.3 Glycosphingolipids

Glycosphingolipids (GSLs) are glycoconjugates that consist of one to several monosaccharides bound to ceramide (Cer). *D. melanogaster* synthesizes 3 types of GSLs: neutral, neutral zwitterionic and acidic zwitterionic (Figure 14). Zwitterionic GSLs have phosphoethanolamine (pEtn) covalently attached to GlcNAc residues and acidic zwitterionic GSLs are terminated with GlcA in addition to pEtn modifications.

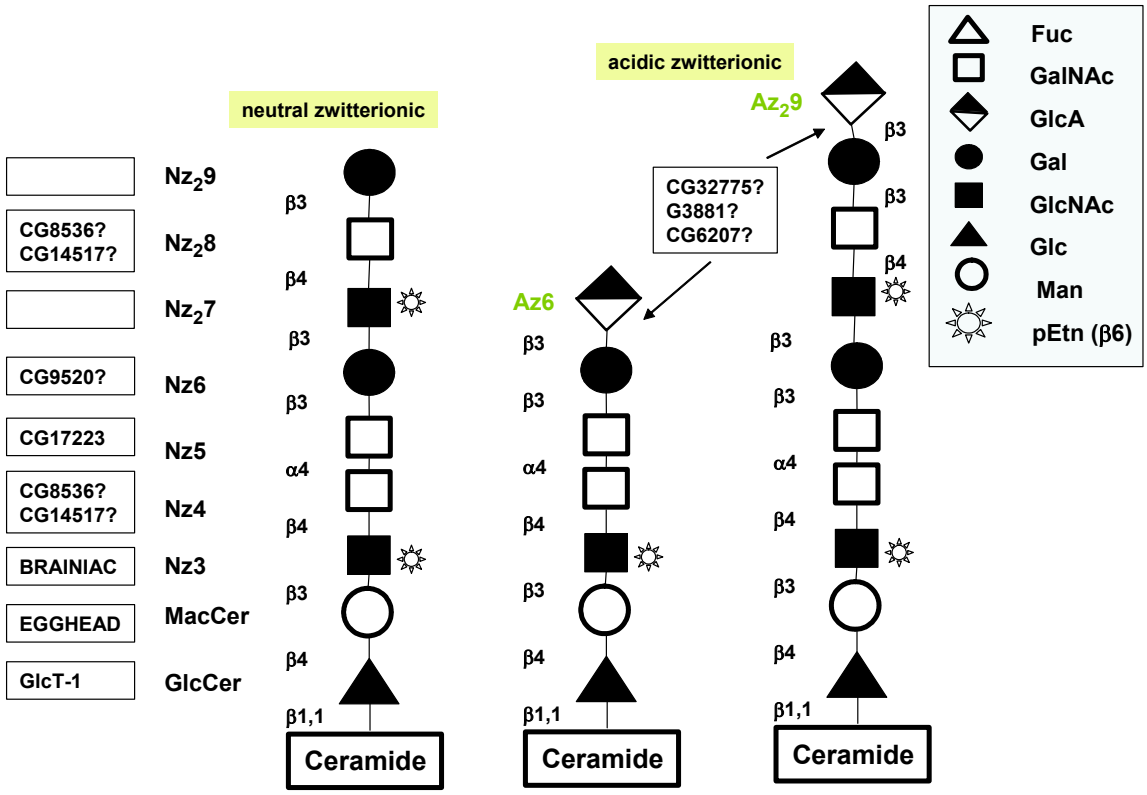

Figure 14: **Biosynthesis of *D. melanogaster* GSLs.** pEtn: phosphoethanolamine, MacCer: Manβ1,4Glc-ceramide. N: neutral, A: acidic (ending with GlcA) z: zwitterionic (carrying pEtn). Enzymes producing the indicated structures are boxed. The question mark denotes sequences with high similarities to characterized enzymes, and/or activities consistent with a possible role in the synthesis of the GSL species that lack pathway assignment. References for individual genes are given in the appendix.

GlcCer is produced by GlcT-1 at the cytosolic face of the ER and is thought to be subsequently flipped into the luminal face of the Golgi by action of an unknown flippase (reviewed in 122) or by an equivalent mechanism. Elongation by the β1,4 mannosyltransferase Egghead (123) yields Manβ1,4Glcβ1,1Cer. The next larger GSL species, N3, is produced through elongation by Brainiac ([I],124). The first zwitterionic GSL is produced from N3 by a pEtn transferase, for which no candidate has been cloned to date, while candidates for the synthesis of the equivalent terminal linkages in Nz4- and $Nz_2$8, CG8536 and CG14517, have been recently cloned and GalNAcT activity was demonstrated on GlcNAcβ with both enzymes (125). *D. melanogaster* Nz5 synthase was cloned and biochemically characterized (126). The action of a β1,3 GalT on Nz5 yields Nz6. Nz6 synthase activity on GSLs catalyzes the same reaction as enzymes that generate the mucin type O-glycan

Galβ1,3GalNAcα, *core1* [III]. Addition of GlcA to Nz6 produces the acidic GSL Az6. Enzymes for the generation of the acidic GSLs Az6 and $Az_2$9 may already be known, since three possible enzymes display GlcAT activity with Galβ terminating acceptors (127). The enzyme catalyzing the elongation of Nz6 to Nz7 with GlcNAc is achieved by a yet unpublished β1,3 GlcNAcT as is the elongation of $Nz_2$8 to $Nz_2$9.

Most of the GSL structures that are present in *D. melanogaster* (82) were identified in GSLs from different dipterans (128, 129), while already lepidopteran insects may have different and/or additional trisaccharylceramide core structures (124). GSLs are ubiquitous components of the animal plasma membrane and more than 400 species of GSLs have been reported, underlining the diversity of this class of glycoconjugates. While invertebrates seem to rely solely on GlcCer as a monosaccharyl-ceramide core (Figure 15), vertebrates also have GalCer. GSLs in vertebrates may be modified with sulfate-groups while nematode GSLs carry phosphatidylcholine groups instead of pEtn. Despite the enormous diversity of mammalian GSLs, mammals seem not to produce Manβ1,4GlcCer (130, [I], [II]). Interestingly, the difference between the core structures Manβ1,4GlcCer and Galβ1,4GlcCer seems to be evolutionary rather than purely functional since substitution of Manβ with Galβ through transgenic expression of a mammalian β1,4GalT complements a lethal mutation in *D. melanogaster egghead* and Galβ1,4GlcCer is elongated (130).

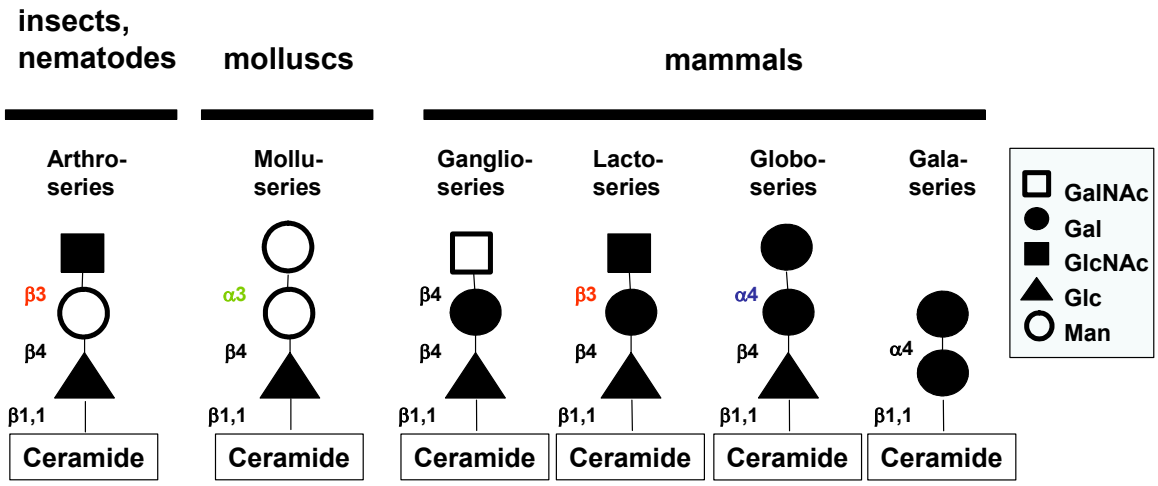

Figure 15: **Evolutionary diversity of GSL core-structures.** Arthro- and mollu-series glycosphingolipids (GSLs) and four types of common mammalian GSLs are shown.

A similarity between vertebrate and arthro-series GSL synthesis is the nature of the terminating monosaccharides. As with GlcA in *D. melanogaster*, mammalian GSLs are often terminally modified by acidic monosaccharides. For instance, Sia in mammalian GSLs is crucial for interactions with lectins like the myelin associated ganglioside receptor (131). In *D. melanogaster* no endogenous lectin binding to GlcA has been reported yet.

While the GSL biosynthetic pathway has been studied extensively in the mouse by reverse genetics, corresponding GTs in invertebrates have not been thoroughly studied, except for the two loci *brainiac* and *egghead* in *D. melanogaster*, which have been genetically characterized without knowledge of their biochemical properties (132-136).

During early embryonic development, *brainiac* and *egghead* mutants display gaps or multiple layers of follicle cells surrounding the oocyte/nurse cell complex (Figure 17). Additionally, *brainiac* alleles also display an overlarge CNS, inspiring the name of the gene (71). Germline clones, eliminating maternal transcripts from *brainiac* (137) and *egghead* (136) revealed that both genes are required in the germline. Both *brainiac* and *egghead* mutants display equivalent and non-additive phenotypes (71) and mutants showed equivalent and strong interactions with mutant alleles of the Transforming Growth Factor α (TGFα)/Epidermal Growth Factor Receptor (EGFR) pathway (137, 138). Therefore, *egghead* and *brainiac wild-type* action was placed into a parallel and cooperative pathway to the EGFR pathway during nurse cell/oocyte – follicle cell interaction (71).

The biochemical data on the interaction between *brainiac* and *egghead* (130) confirmed those obtained by the genetic approach, both gene products acting in the same GSL synthetic pathway. On the other hand, the nature and the level of interaction(s) between GSL synthetic genes with components of the EGFR pathway remain unresolved to date.

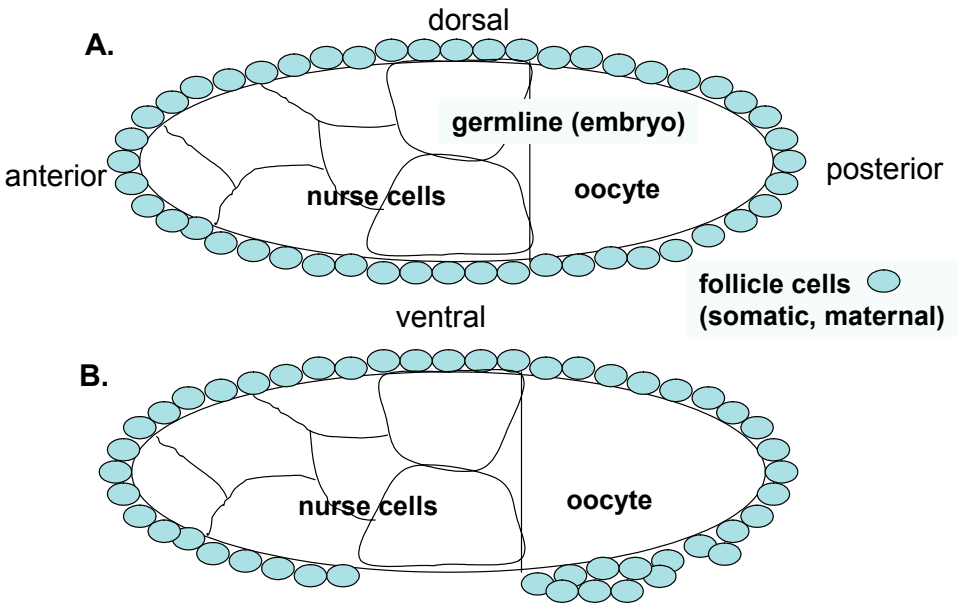

Figure 16: The layer of follicle cells surrounding oocyte/nurse cell is affected by mutations at the GSL synthetic loci *brainiac* and *egghead*. A. representation of a wild-type nurse cell/oocyte – follicle cell monolayer complex. **B.** *brainiac* and *egghead* display gaps and multiple layers of follicle cells around the complex.
Both transcripts are necessary in the germline (oocyte/nurse cells) and *brainiac* mutant embryos later display a fused dorsal appendage, which indicates defects in the specification of the dorsoventral axis of the chorion. Dorsoventral polarity defects can also be observed in alleles of the EGFR (necessary in the follicle cells) and its signal, encoded by the TGFα homologous gene *gurken* (required in the germline) (139, 140). However, *EGFR* and *gurken* alleles also display polarity defects in the embryo *proper*. Double mutants of *brainiac* or *egghead* with both *EGFR* and *gurken* alleles show stronger phenotypes than null alleles of *EGFR* or *gurken* mutants.

8.4 O-Fucose initiated O-glycosylation

Proteins, that are O-fucosylated typically contain EGFR repeats with the consensus sequence CX_2G_2S/TC (X is any amino acid and S/T the O-Fuc site) between the second and third conserved cysteine (141) or they contain thrombospondin 1 (TSR) repeats with the sequence $WX_5CX_{2/3}S/TCX_2G$ (52). Both TSR repeats and EGFR repeats are common modules on animal cell surface proteins.

EGFR repeats are 30-40 amino acids long and contain 6 cysteins that form 3 disulfide bridges, maintaining the structure of the module. EGFR repeats are found on receptors like EGFR, Delta, Serrate (Serrate orthologues are called "Jagged" in vertebrates) and Notch. While mammals encode two O-FucTs, only one gene encoding an O-FucT enzyme, *neurotic*, has been described in *D. melanogaster* to date (80, 142, 143). A β1,3 GlcNAcT encoded by *fringe* subsequently catalyzes the formation of the disaccharide GlcNAcβ1,3Fucα (15) on a subset of O-fucosylated EGFR repeats (Figure 17). In mammals, the disaccharide is sometimes elongated by β1,4 GalT-I and an α2,3 SiaT (51). However, there is no evidence yet that EGFR modules of *D. melanogaster* are elongated any further than with GlcNAc.

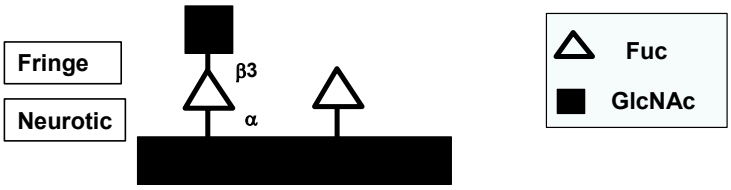

Figure 17: EGFR type O-Fucose glycosylation. Neurotic and Fringe catalyze the formation of the disaccharide GlcNAcβ1,3Fucα-Ser/Thr on EGFR modules in *D. melanogaster*.

The Notch protein contains 36 EGFR repeats and mediates a plethora of cell fate decisions during *D. melanogaster* development. For example, both positive and negative signals from the Notch signalling pathway regulate the differentiation of the central and peripheral nervous systems, the eyes, antennae and legs. Oogenesis and wing margin formation (Figure 18) are also systems that are influenced by Notch (144). Flies exhibiting notching at their wing edges inspired the name of *Notch* mutants. The ligands for the Notch receptor, Delta and Serrate interact with EGFR repeats on Notch, and O-Fuc initiated glycans on EGFR repeats modifies the affinity of the interactions between Notch and its ligands. The modulation of this receptor-ligand interaction made Fringe the best studied GT in terms of genetic analysis and many reviews have addressed Fringe in connection with its substrate, the Notch receptor (145-150).

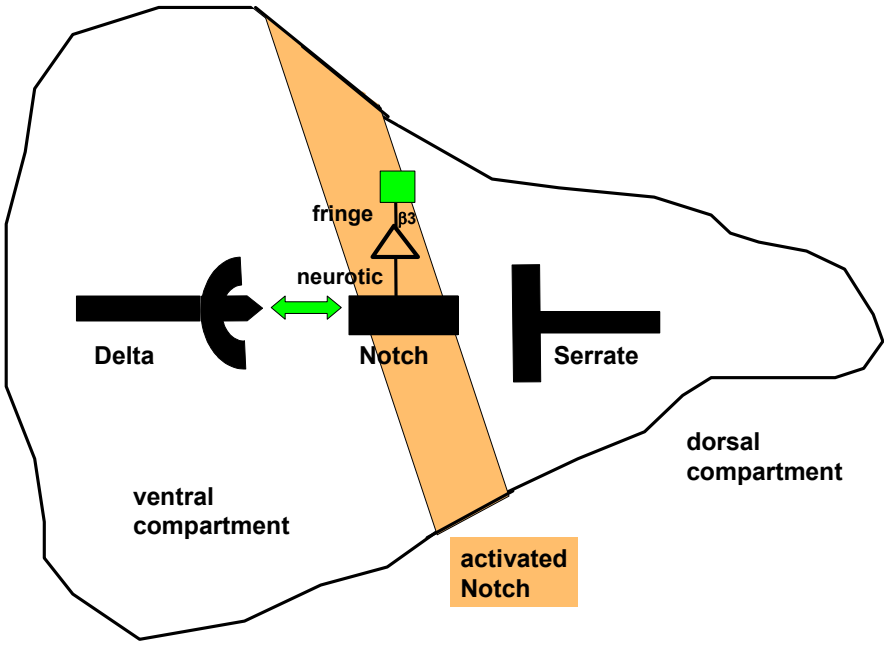

Figure 18: Fringe modifies Notch-Delta/Notch-Serrate interaction in the wing disc. A thin stripe of activated Notch (i.e. Wingless producing) cells in the middle of the wing disc demarcates the border between its dorsal and its ventral portion (151), the border being the place where the wing blade is situated in adults. Although the *Notch* gene is initially expressed throughout the wing disc, it is only activated at its dorsal-ventral boundary upon interaction with two groups of cells, one group carrying Serrate and the other Delta on their surface (147, 152). Fringe is expressed in the dorsal compartment of the wing disc only, where *apterous* drives its expression (153, 154) and the Fringe product inhibits Serrate-Notch interactions in the dorsal compartment. Delta and Serrate levels are regulated across the dorsal- ventral border of the wing disc in a positive feedback loop and the Fringe product on Notch potentates Notch-Delta binding on the dorsal-ventral border (155). The β1,3 GlcNAcT Fringe is therefore necessary for wing blade formation.

Clonal analysis in wing discs revealed internal ectopic wing margins at the border of adjacent *fringe*$^{+}$ and *fringe*$^{-}$ clones (70), pointing towards a role of *fringe* in *Notch* mediated wing border formation. Genetic analysis showed that *fringe* acts cell-autonomously and in signal receiving cells (155). This result is consistent with biochemical data (15) indicating that only the Notch receptor and not its ligands Delta and Serrate, which also carry EGFR modules, are the substrate for the β1,3 GlcNAcT Fringe. The consequence of the modification of Notch-EGFR repeats by Fringe is that the Notch receptor binds much more to Delta and less to Serrate than without GlcNAc modification when assayed *in vitro* (15).

The *fringe* phenotypes observed are not as strong as phenotypes of its biochemical precursor *neurotic*, which also impacts cell fate decisions in the developing CNS (156). This might indicate that *fringe* is only involved in processes during which the Notch receptor requires to distinguish between Delta and Serrate (reviewed in ref. 157). Whether this is the case in all developmental processes affected by *fringe* remains to be assessed.

8.5 Mucin Type O-Glycosylation

Mucins are proteins that carry GalNAcα-O-Ser/Thr initiated glycans. Typically, mucin-type glycans are found on protein domains rich in Thr, Ser and Pro. Their synthesis is initiated by polypeptide α–*N*-acetylgalactosaminyltransferases (ppGalNAcTs) in the *cis*-Golgi (158). This first step results in the generation of the Tn-antigen (159), representing GalNAcα-O-Ser/Thr. Nine active *D. melanogaster* ppGalNAcTs have been cloned and characterized with respect to their activities to date (160-162).

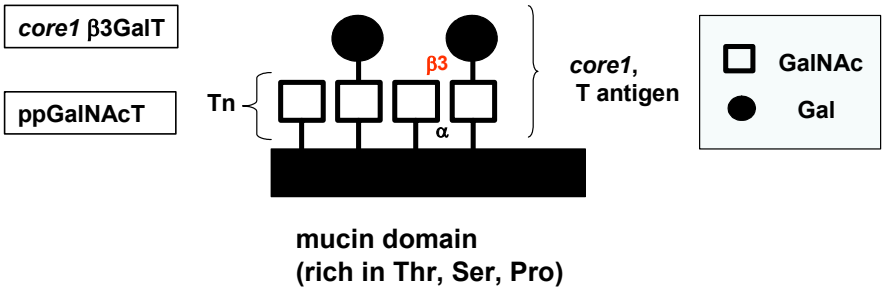

Figure 19. **Mucin-type *O*-glycosylation in *D. melanogaster*** has not yet been shown to extend the *core1*-structure, i.e. Galβ1,3GalNAcα Ser/Thr. ppGalNAcTs act on mucin domains within proteins or are glycopeptide specific (mucins with GalNAc-O-Ser/Thr at an adjacent site) to produce the Tn antigen. Tn is extended to the T-antigen or *core1* by *core1* β1,3 GalTs.

The family of ppGalNAcT enzymes can be categorized into activities that glycosylate peptides and activities that act on glycopeptides, carrying a Tn-antigen at an adjacent site. A subclassification of activities is achieved by their sequence-specificity. One reason for the high number of ppGalNAcT encoding genes may be the structural diversity of their substrates. While highly complex mucin type *O*-glycans have been reported for *C. elegans* (163), *D. melanogaster* mucin type *O*-glycans have not been established beyond the *core1* structure Galβ1,3GalNAcα -O-Thr/Ser. However, the limited number of structural studies (164, 165) precludes any conclusion regarding *core1* elongation in *D. melanogaster*. Four *core1* β1,3 GalT enzymes catalyzing the elongation of GalNAcα with Gal have been cloned and characterized [III]. In mammals, only one *core1* β1,3 GalT has been cloned to date (166). Mammals, however, present a previously unexpected factor in the regulation of mucin-type glycosylation. A *core1* β1,3 GalT- specific chaperone, *cosmc* (*core1 specific molecular chaperone*) (167) prevents the GT from proteasomal degradation. Interestingly, the primary sequence of *cosmc* shares many features of the *core1* β1,3 GalT itself (168).

Mucins provide a protective barrier and lubrication at the apical surface of epithelial cells (169) and are under discussion to provide targeting information as for example to the apical brush border membrane (170). In *D. melanogaster*, six mucin-type salivary gland secretion proteins (glue proteins) are found in large quantities in the salivary glands of third instar larvae (171), and salivary gland mucins were shown to be important for the entrapment and elimination of bacteria (172). Occurrence of the *core1* structure in *D. melanogaster* has also been reported on haemocytes (173), on intercellular bridges between germline and somatic cells (174) and on several peptides involved in innate immunity (reviewed in refs. 175 and 176). Moreover, mucin type O-glycans on the anti-microbial peptide Drosocin have been demonstrated to increase its potency (164). In the latter case, however, a NMR study of glycosylated versus unglycosylated Drosocin has shown that the glycan itself is not implicated in the binding to microorganisms (177).

The scarcity of data on mutants in mucin type O-glycosylation disallows to draw conclusions on the importance or function of these glycans in *D. melanogaster* to date. However, one *D. melanogaster* ppGalNAcT, encoded by *Pgant35A* (161, 178), was demonstrated to be essential and a deletion encompassing the *Pgant35A* locus leads to death during pupation (179). In humans, a mutation in a ppGalNAcT leads to tumoral calcinosis, a metabolic disorder characterized by massive calcium deposits in the skin and subcutaneous tissues (180). However, since both acceptor specificity and expression patterns of some ppGalNAcTs overlap, rescue with other ppGalNAcTs likely accounts for the lack of a visible phenotype for several ppGalNAcT knock-out models in mice (181, 182).

While no *D. melanogaster* mutant has been described for the formation of *core1*, a mouse knock-out for the single *core1* β3GalT demonstrated the importance of that activity in mammals since early embryonic lethality was observed (183). The *core1* β3GalT1-null mice exhibited defects in angiogenesis and haemorrhages possibly caused by defective interactions between endothelial cells and the extracellular matrix.

9 β1,3 Glycosyltransferases of D. melanogaster

As outlined earlier (section 7), the *in vitro* characterization of GTs is the first step toward understanding the functions of glycan chains.

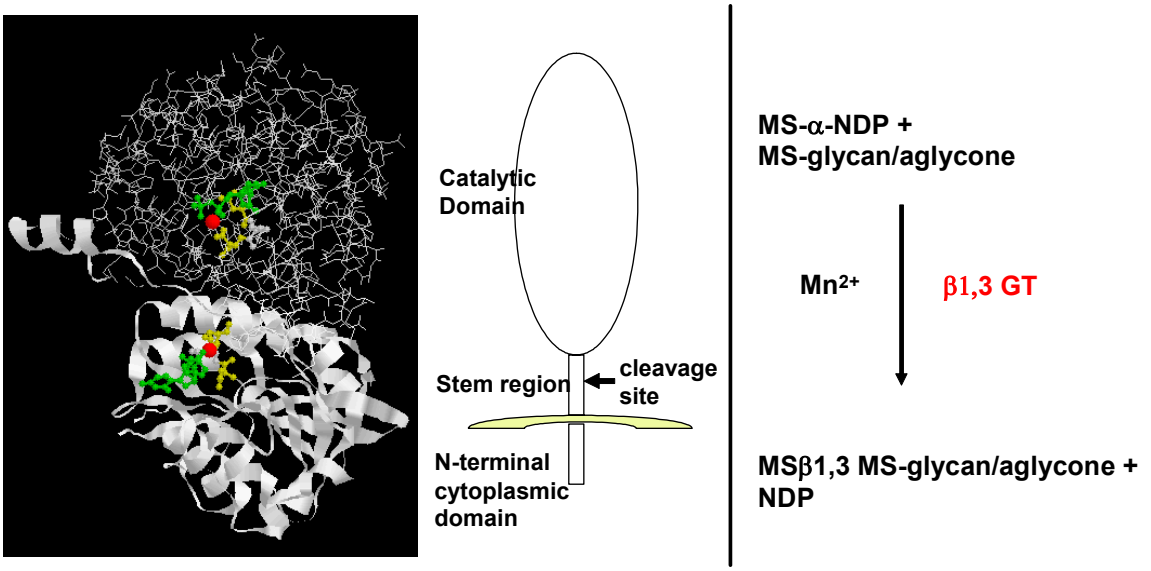

Figure 20: Structure of and reaction catalyzed by β1,3 GTs. Left: human GlcAT-I dimer (46, PDB:1KWS) in complex with UDP-GlcA (green) and Mn^{2+} (red). DXD motif (c.f. Table 6) in yellow. One monomer is drawn in ribbons, the other in backbone representation. **Middle:**. β1,3 GTs are type II trans- membrane proteins with a single transmembrane domain anchoring the protein, a short N-terminal domain and a longer C-terminal domain. The short N-terminus is located within the cytosol while the C-terminus, which includes the catalytic domain, lies in the lumen of the Golgi compartment. β1,3 linkages in *D. melanogaster* glycans are found in GAGs (section 8.2), GSLs (section 8.3), O-Fuc initiated O-Glycans (section 8.4) and mucin-type O-glycans (section 8.5). **Right:** Reaction catalyzed by β1,3 GTs. β1,3 GTs often require a pH between 5 and 7 and the divalent cation Mn^{2+} for activity (184). The catalytically active part of the transferase defines the donor and acceptor specificity of the transferase. With most β1,3GTs, the acceptor monosaccharide and its anomericity provides sufficient specificity for activity measurements, although a penultimate monosaccharide corresponding to the physiological substrate often increases the measured activity.

Sequence alignments of characterized *D. melanogaster* GTs reveal clustering according to the linkage catalyzed and to some degree also by donor specificity (Figure 21). Along this line, the two *D. melanogaster* genes, *brainiac* (185) and *fringe* (70) were predicted to encode β1,3 GTs (186) by virtue of their sequence similarities to known bacterial enzymes. Similarly, owing to their cross-species homologies, as many as 21 candidate β1,3 GTs (Figure 22) can be retrieved using BLAST (60) searches.

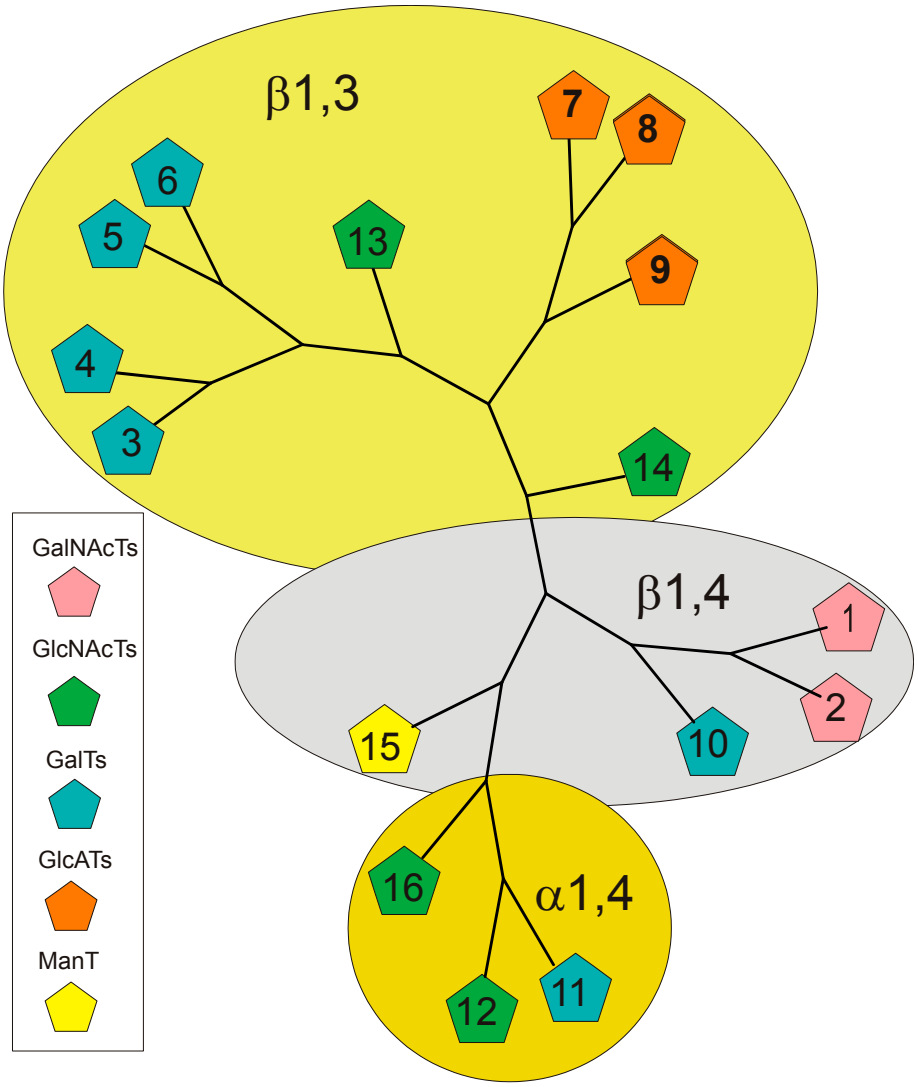

Figure 21: Unrooted ClustalW (63) **tree of characterized *D. melanogaster* GTs** with different donor and acceptor specificities. Sequence similarity among GTs is more related to the linkage catalyzed than to the donor specifity of the enzyme. For references see appendix.
Enzymes: 1: CG14517, 2: CG8536, 3: CG9520, 4: CG8708, 5:CG13904-1, 6: CG2975, 7: CG3881, 8: CG6207, 9: CG32775, 10: CG11780, 11: BOTV, 12: TTV (β4GlcNAcT, β3GlcAT, co-polymerase), 13: BRN, 14: FRG, 15: EGH, 16: SOTV (β4GlcNAcT, β3GlcAT, co-polymerase).

Among the most conserved areas are four short motifs (Table 6). Mutations in motif I (TW), motif II (DXD) and motif IV (EDV) reduced or abrogated activity of a murine β1,3 GT (187). One would expect conserved β1,3 GTs to retain their donor and acceptor specificity, however, global sequence comparisons reveal that the degree of cross-species sequence similarity of orthologous β1,3 GTs may be relatively small, as with the Manβ specific β1,3 GlcNAcTs Bre-5 and Brainiac (number 27 and 31 in Figure 23), or high, as exemplified by the Fringe family of orthologues (numbers 15 to 18 in Figure 23). Furthermore, protein mutant analysis for β1,3 GlcAT-I showed, that the determinants for its donor specificity were located to mostly non-conserved regions close to the C-terminus of the catalytic domain (47). For example, the substitution of a single amino acid at the C-terminus broadened the donor specificity of the enzyme considerably from UDP-GlcA to include the nucleotide sugars UDP-Glc, UDP-GlcNAc and (non-physiological) UDP-Man. It is therefore difficult to predict, which substrates are preferred and the determination of the specificities of a GT remains “wet” biochemistry.

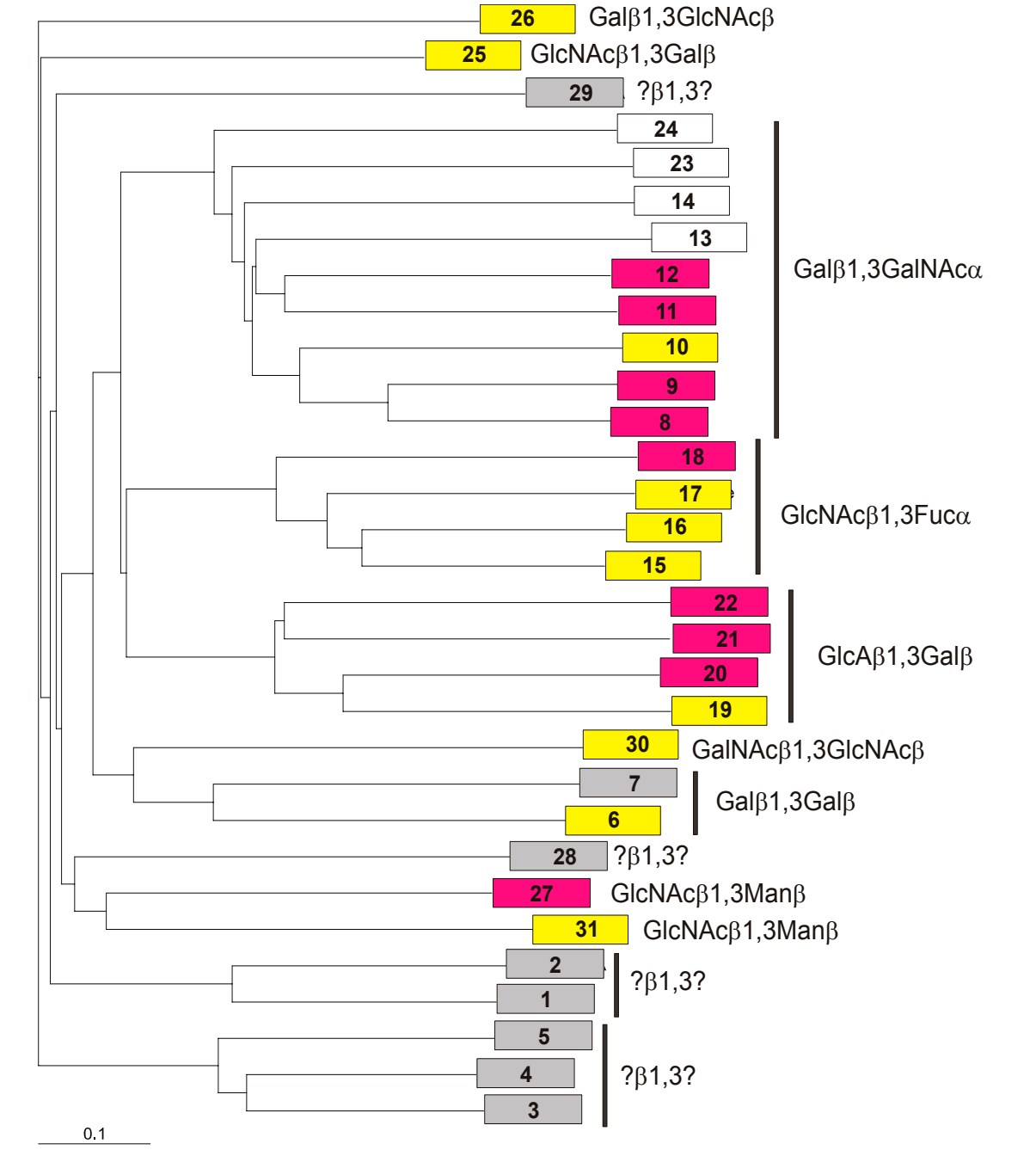

Figure 22: **ClustalW Tree (62) with sequences of putative and established β1,3 GTs** of *D. melanogaster* and other species. yellow: non-*D. melanogaster* GTs with established donor and acceptor specificities, red: *D. melanogaster* GTs with established specificities, white: expressed isoforms w/o assigned activities [III], grey: unpublished. **Sequences** (accession number): **1**: CG8668, **2**: CG8673, **3**: CG30036, **4**: CG33145, **5**: CG33037, **6**: hβ3GalT6 (NP_542172), **7**: CG8734, **8**: CG9520, **9**: CG8708, **10**: hC1β3GalT (AAF81981), **11**: CG13904-1, **12**: CG2975, **13**: CG2983, **14**: CG7440, **15**: rradical fringe (Q9R149), **16**: clunatic fringe (O12971), **17**: zfmanic fringe (AAT46070), **18**: FRG, **19**: hGlcAT-1 (O94766), **20**: CG32775, **21**: CG3881, **22**: CG6207, **23**: CG18558, **24**: CG3119, **25**: mβ3GnT5 (NP_114436), **26**: hβ3GalT5 (Q9J167), **27**: BRN, **28**: CG11357, **29**: CG3038, **30**: mβ3GalNAcT-II (BC029564), **31**: celbre-5 (NP_741492). Prefixes as follows: zf: zebra fish, m: mouse, r: rat, h: human, c: chicken, cel: *C. elegans*, for *D. melanogaster* sequences c.f. Table 6 on the next page.

Table 6: Motifs conserved in established and putative *D. melanogaster* β1,3 Glycosyltransferases

Gene[REF]	Motif I	Motif II	Motif III	Motif IV	Activity
CG3881-RB[1]	TPTYPR	YFGDDDNTY	YGVSGPV	EDLFLRS	GlcAβ1,3Galβ
CG6207-RA[1]	TPTYRR	YFADDDNTY	VTKTGVS	EDGFLRS	GlcAβ1,3Galβ
CG32775-RA[1]	TPTYPR	FFMDDDNSY	VGLVGGL	TDVLVWH	GlcAβ1,3Galβ
CG9520[2]	KRTWGK	LKADDDTYT	GYMSGGA	EDVEIGK	Galβ1,3GalNAcα
CG8708[2]	KRTWGK	YKADDDTYA	GYMSGGA	EDIEIGR	Galβ1,3GalNAcα
CG13904-1[2]	KGTWGR	LKADDDTFV	GYMSGGA	EDRQIGF	Galβ1,3GalNAcα
CG2975[2]	KRTWGR	LKADDDTYF	GYMSGGA	EDVELGR	Galβ1,3GalNAcα
CG18558[2]	SETWGR	YKADDDTYA	LYMSGGA	EDYVMGK	n.d.
CG3119[2]	YETWGQ	LKADDDTYV	SYMSGGA	EDFYMGI	n.d.
CG2983[2]	LETWGR	LEADDETYV	n.p.	n.p.	n.d.
CG7440[2]	LRTWGK	IKADDDTLL	SYMSGGS	EDVEMGK	n.d.
FRG[3]	IKTWFQ	CHFDDDNYV	WFATGGA	DDVTMGF	GlcNAcβ1,3Fucα
CG33145-RB	RETWGN	LKCDDDTFV	KYLSGAG	EDVYITG	n.d.
CG30036-RA	RQTWGN	LKCDDDTFV	KYLSGTG	EDVFVTG	n.d.
CG30037-RA	RKLWGN	FKCDDDTFV	RYLCGSG	EDMFVTG	n.d.
CG8734	n.p	LKVDDDTYV	YALGGGY	EDVSVGT	n.d.
CG8668	RQTWMH	LKTDDDMFI	SFTTGPA	EDVFTTG	n.d.
CG8673-RA	RQTWMH	LKTDDDMFI	YFTTGPA	EDVFTTG	n.d.
BRN[4]	RRTWGY	LFVDDDYYV	PYVTAGA	DDVYLGI	GlcNAcβ1,3Manβ
CG3038-RA	n.p.	IKLDDDIIY	AYLSGWL	DDTWLTG	n.d.
CG11357	RETWAN	VKVDDDVFM	EYCPGMA	DDVLITG	n.d.

Individual motifs were aligned using the local alignment program LFASTA (188) after TBLASTN (60) searches using *D. melanogaster* amino acid sequences from Fringe and Brainiac as well as established mammalian, *C. elegans* and zebra fish GTs on www.flybase.org identified 21 expressed transcripts (c.f. Figure 22). If differences between amino acid sequences of predicted transcripts were present, the gene most homologous to Brainiac was chosen.

n.p.: not present, n.d.: not determined. References: [1](92), [2][III], [3](15), [4]([I], [II], and 124)

10 Abbreviations

GT: glycosyltransferase, Gal: Galactose, GalT: galactosyltransferase, GlcNAc: N-acetylglucosamine, GlcNAcT: N-acetylglucosaminyltransferase, GalNAc: N-acetylgalactosamine, Fuc: Fucose, Glc: Glucose, IdoA: Iduronic acid, GlcA: Glucuronic acid, Dolichylphosphomannose: DolP-Man, Dolichylphosphoglucose: DolP-Glc, DolPP: dolichylpyrophosphate, GAG: glycosaminoglycan, CS: chondroitin sulfate, HS: heparan sulfate, GSL: glycosphingolipid, Cer: Ceramide, GlcCer: Glucosylceramide, MacCer: Mactosylceramide, pEtn: phosphoethanolamine, PC: phosphocholine, pEtnT: phosphoethanolamine transferase, GPI-anchor: glycosylphosphatidylinositol anchors, Nz: neutral zwitterionic (contains 1 pEtn residue), Nz_2: neutral doubly zwitterionic (contains 2 pEtn residues), Az: acidic zwitterionic, EGFR epidermal growth factor receptor, TSP: thrombospondin 1, ppGalNAcT: polypeptide N-acetylgalactosaminyltransferases.

11 Literature cited

[I] Muller, R., Altmann, F., Zhou, D., and Hennet, T. (2002). The *Drosophila melanogaster brainiac* protein is a glycolipid-specific beta 1,3N-acetylglucosaminyltransferase. J Biol Chem *277*, 32417-32420.

[II] Griffitts, J. S., Huffman, D. L., Whitacre, J. L., Barrows, B. D., Marroquin, L. D., Muller, R., Brown, J. R., Hennet, T., Esko, J. D., and Aroian, R. V. (2003). Resistance to a bacterial toxin is mediated by removal of a conserved glycosylation pathway required for toxin-host interactions. J Biol Chem *278*, 45594-45602.

[III] Muller, R., Hulsmeier, A. J., Altmann, F., Ten Hagen, K., Tiemeyer, M., and Hennet, T. (2005). Characterization of mucin-type core-1 beta1-3 galactosyltransferase homologous enzymes in Drosophila melanogaster. Febs J *272*, 4295-4305.

1. Keppler OT, Horstkorte R, Pawlita M, Schmidt C, Reutter W. 2001. *Glycobiology* 11: 11R-8R
2. Laine RA. 1994. *Glycobiology* 4: 759-67
3. Hennet T. 2002. *Cell Mol Life Sci* 59: 1081-95
4. Ferrero-Garcia MA, Trombetta SE, Sanchez DO, Reglero A, Frasch AC, Parodi AJ. 1993. *Eur J Biochem* 213: 765-71
5. Spiro RG. 2002. *Glycobiology* 12: 43R-56R
6. Varki A. 1999. *(Ed.) Essentials of Glycobiology, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press*
7. Trombetta ES, Helenius A. 1998. *Curr Opin Struct Biol* 8: 587-92
8. Mehta A, Lu X, Block TM, Blumberg BS, Dwek RA. 1997. *Proc Natl Acad Sci U S A* 94: 1822-7
9. Alper J. 2001. *Science* 291: 2338-43
10. Wang P, Granados RR. 1997. *Proc Natl Acad Sci U S A* 94: 6977-82
11. Huet G, Gouyer V, Delacour D, Richet C, Zanetta JP, et al. 2003. *Biochimie* 85: 323-30
12. Parodi AJ. 2000. *Biochem J* 348 Pt 1: 1-13
13. Brady RO, Murray GJ, Barton NW. 1994. *J Inherit Metab Dis* 17: 510-9
14. Galili U, Rachmilewitz EA, Peleg A, Flechner I. 1984. *J Exp Med* 160: 1519-31
15. Bruckner K, Perez L, Clausen H, Cohen S. 2000. *Nature* 406: 411-5

16. Miljan EA, Bremer EG. 2002. *Sci STKE* 2002 RE15: 160: 1-10
17. Schaaf L, Leiprecht A, Saji M, Hubner U, Usadel KH, Kohn LD. 1997. *Mol Cell Endocrinol* 132: 185-94
18. Slawson C, Hart GW. 2003. *Curr Opin Struct Biol* 13: 631-6
19. Takeuchi M, Inoue N, Strickland TW, Kubota M, Wada M, et al. 1989. *Proc Natl Acad Sci U S A* 86: 7819-22
20. Baenziger JU, Kumar S, Brodbeck RM, Smith PL, Beranek MC. 1992. *Proc Natl Acad Sci U S A* 89: 334-8
21. Bell GI, Burant CF, Takeda J, Gould GW. 1993. *J Biol Chem* 268: 19161-4
22. Loo DD, Hirayama BA, Gallardo EM, Lam JT, Turk E, Wright EM. 1998. *Proc Natl Acad Sci U S A* 95: 7789-94
23. Raas-Rothschild A, Pankova-Kholmyansky I, Kacher Y, Futerman AH. 2004. *Glycoconj J* 21: 295-304
24. Aebi M, Hennet T. 2001. *Trends Cell Biol* 11: 136-41
25. Perez M, Hirschberg CB. 1986. *J Biol Chem* 261: 6822-30
26. Sommers LW, Hirschberg CB. 1982. *J Biol Chem* 257: 10811-7
27. Waldman BC, Rudnick G. 1990. *Biochemistry* 29: 44-52
28. Hirschberg CB, Robbins PW, Abeijon C. 1998. *Annu Rev Biochem* 67: 49-69
29. Kapitonov D, Yu RK. 1999. *Glycobiology* 9: 961-78
30. Brew K, Vanaman TC, Hill RL. 1968. *Proc Natl Acad Sci U S A* 59: 491-7
31. Lind T, Tufaro F, McCormick C, Lindahl U, Lidholt K. 1998. *J Biol Chem* 273: 26265-8
32. Roseman S. 1970. *Chem Phys Lipids* 5: 270-97
33. Opat AS, van Vliet C, Gleeson PA. 2001. *Biochimie* 83: 763-73
34. Varki A, and Marth, J. 1995. *Seminars in Developmental Biology* 6: 127-38
35. Berger EG. 2002. *Glycobiology* 12: 29R-36R
36. Borsig L, Katopodis AG, Bowen BR, Berger EG. 1998. *Glycobiology* 8: 259-68
37. Homa FL, Hollander T, Lehman DJ, Thomsen DR, Elhammer AP. 1993. *J Biol Chem* 268: 12609-16
38. Munro S, Freeman M. 2000. *Curr Biol* 10: 813-20
39. Wu X, Steet RA, Bohorov O, Bakker J, Newell J, et al. 2004. *Nat Med* 10: 518-23
40. Bretscher MS, Munro S. 1993. *Science* 261: 1280-1
41. Nilsson T, Slusarewicz P, Hoe MH, Warren G. 1993. *FEBS Lett* 330: 1-4
42. Ju T, Cummings RD, Canfield WM. 2002. *J Biol Chem* 277: 169-77

43. Correia T, Papayannopoulos V, Panin V, Woronoff P, Jiang J, et al. 2003. *Proc Natl Acad Sci U S A* 100: 6404-9
44. McCormick C, Duncan G, Goutsos KT, Tufaro F. 2000. *Proc Natl Acad Sci U S A* 97: 668-73
45. Giraudo CG, Daniotti JL, Maccioni HJ. 2001. *Proc Natl Acad Sci U S A* 98: 1625-30
46. Pedersen LC, Darden TA, Negishi M. 2002. *J Biol Chem* 277: 21869-73
47. Ouzzine M, Gulberti S, Levoin N, Netter P, Magdalou J, Fournel-Gigleux S. 2002. *J Biol Chem* 277: 25439-45
48. Ten Hagen KG, Fritz TA, Tabak LA. 2003. *Glycobiology* 13: 1R-16R
49. Coste, H., Martel, M. B., and Got, R. (1986) *Biochim Biophys Acta* **858**, 6-12
50. Jeckel D, Karrenbauer A, Burger KN, van Meer G, Wieland F. 1992. *J Cell Biol* 117: 259-67
51. Moloney DJ, Shair LH, Lu FM, Xia J, Locke R, et al. 2000. *J Biol Chem* 275: 9604-11
52. Gonzalez de Peredo A, Klein D, Macek B, Hess D, Peter-Katalinic J, Hofsteenge J. 2002. *Mol Cell Proteomics* 1: 11-8
53. Luo Y, Haltiwanger RS. 2005. *J Biol Chem* 280: 11289-94
54. Strahl-Bolsinger S, Gentzsch M, Tanner W. 1999. *Biochim Biophys Acta* 1426: 297-307
55. Wilson IB. 2004. *Cell Mol Life Sci* 61: 794-809
56. Prydz K, Dalen KT. 2000. *J Cell Sci* 113 Pt 2: 193-205
57. Krieg J, Glasner W, Vicentini A, Doucey MA, Loffler A, et al. 1997. *J Biol Chem* 272: 26687-92
58. Eisenhaber B, Maurer-Stroh S, Novatchkova M, Schneider G, Eisenhaber F. 2003. *Bioessays* 25: 367-85
59. Datta AK, Paulson JC. 1997. *Indian J Biochem Biophys* 34: 157-65
60. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. *J Mol Biol* 215: 403-10
61. Ramakrishnan B, Qasba PK. 2002. *J Biol Chem* 277: 20833-9
62. Felsenstein J. 1989. *Cladistics* 5: 164-6.
63. Thompson JD, Higgins DG, Gibson TJ. 1994. *Nucleic Acids Res* 22: 4673-80
64. Grunewald S, Matthijs G, Jaeken J. 2002. *Pediatr Res* 52: 618-24
65. Wildt S, Gerngross TU. 2005. *Nat Rev Microbiol* 3: 119-28
66. Stanley P, Caillibot V, Siminovitch L. 1975. *Cell* 6: 121-8
67. Lowe JB, Marth JD. 2003. *Annu Rev Biochem* 72: 643-91
68. Herman T, Hartwieg E, Horvitz HR. 1999. *Proc Natl Acad Sci U S A* 96: 968-73
69. Herman T, Horvitz HR. 1999. *Proc Natl Acad Sci U S A* 96: 974-9

70. Irvine KD, Wieschaus E. 1994. *Cell* 79: 595-606
71. Goode S, Melnick M, Chou TB, Perrimon N. 1996. *Development* 122: 3863-79
72. Smith JE, 3rd, Cummings CA, Cronmiller C. 2002. *Development* 129: 3255-67
73. Selva EM, Hong K, Baeg GH, Beverley SM, Turco SJ, et al. 2001. *Nat Cell Biol* 3: 809-15
74. Katz F, Moats W, Jan YN. 1988. *Embo J* 7: 3471-7
75. Seppo A, Tiemeyer M. 2000. *Glycobiology* 10: 751-60
76. Wilson IB. 2002. *Curr Opin Struct Biol* 12: 569-77
77. Solter D, Knowles BB. 1978. *Proc Natl Acad Sci U S A* 75: 5565-9
78. Parker GF, Williams PJ, Butters TD, Roberts DB. 1991. *FEBS Lett* 290: 58-60
79. Fabini G, Freilinger A, Altmann F, Wilson IB. 2001. *J Biol Chem* 276: 28058-67
80. Moloney DJ, Panin VM, Johnston SH, Chen J, Shao L, et al. 2000. *Nature* 406: 369-75
81. Martin-Blanco E, Garcia-Bellido A. 1996. *Proc Natl Acad Sci U S A* 93: 6048-52
82. Seppo A, Moreland M, Schweingruber H, Tiemeyer M. 2000. *Eur J Biochem* 267: 3549-58
83. Krieg J, Hartmann S, Vicentini A, Glasner W, Hess D, Hofsteenge J. 1998. *Mol Biol Cell* 9: 301-9
84. Hortsch M, Goodman CS. 1990. *J Biol Chem* 265: 15104-9
85. Spring J, Paine-Saunders SE, Hynes RO, Bernfield M. 1994. *Proc Natl Acad Sci U S A* 91: 3334-8
86. Kelly WG, Hart GW. 1989. *Cell* 57: 243-51
87. Helenius J, Ng DT, Marolda CL, Walter P, Valvano MA, Aebi M. 2002. *Nature* 415: 447-50
88. Kleizen B, Braakman I. 2004. *Curr Opin Cell Biol* 16: 343-9
89. Katanosaka K, Tokunaga F, Kawamura S, Ozaki K. 1998. *FEBS Lett* 424: 149-54
90. Kurzik-Dumke U, Kaymer M, Gundacker D, Debes A, Labitzke K. 1997. *Gene* 200: 45-58
91. Korner C, Knauer R, Stephani U, Marquardt T, Lehle L, von Figura K. 1999. *Embo J* 18: 6816-22
92. Takeo S, Fujise M, Akiyama T, Habuchi H, Itano N, et al. 2004. *J Biol Chem* 279: 18920-5
93. Pinto DO, Ferreira PL, Andrade LR, Petrs-Silva H, Linden R, et al. 2004. *Glycobiology* 14: 529-36
94. Toyoda H, Kinoshita-Toyoda A, Fox B, Selleck SB. 2000. *J Biol Chem* 275: 21856-61

95. Mizuguchi S, Uyama T, Kitagawa H, Nomura KH, Dejima K, et al. 2003. *Nature* 423: 443-8
96. Wilson IB. 2002. *J Biol Chem* 277: 21207-12
97. Han C, Belenkaya TY, Khodoun M, Tauchi M, Lin X. 2004. *Development* 131: 1563-75
98. Bornemann DJ, Duncan JE, Staatz W, Selleck S, Warrior R. 2004. *Development* 131: 1927-38
99. Perrimon N, Lanjuin A, Arnold C, Noll E. 1996. *Genetics* 144: 1681-92
100. Anderson KV, Jurgens G, Nusslein-Volhard C. 1985. *Cell* 42: 779-89
101. Sergeev P, Streit A, Heller A, Steinmann-Zwicky M. 2001. *Dev Dyn* 220: 122-32
102. Kreuger J, Salmivirta M, Sturiale L, Gimenez-Gallego G, Lindahl U. 2001. *J Biol Chem* 276: 30744-52
103. Lortat-Jacob H, Turnbull JE, Grimaud JA. 1995. *Biochem J* 310 (Pt 2): 497-505
104. Yoneda A, Asada M, Oda Y, Suzuki M, Imamura T. 2000. *Nat Biotechnol* 18: 641-4
105. Reichsman F, Smith L, Cumberledge S. 1996. *J Cell Biol* 135: 819-27
106. Moline MM, Dierick HA, Southern C, Bejsovec A. 2000. *Curr Biol* 10: 1127-30
107. Lin X, Perrimon N. 1999. *Nature* 400: 281-4
108. Jackson SM, Nakato H, Sugiura M, Jannuzi A, Oakes R, et al. 1997. *Development* 124: 4113-20
109. Park Y, Rangel C, Reynolds MM, Caldwell MC, Johns M, et al. 2003. *Dev Biol* 253: 247-57
110. Voigt A, Pflanz R, Schafer U, Jackle H. 2002. *Dev Dyn* 224: 403-12
111. Cadigan KM. 2002. *Semin Cell Dev Biol* 13: 83-90
112. Neumann C, Cohen S. 1997. *Bioessays* 19: 721-9
113. Willert K, Brown JD, Danenberg E, Duncan AW, Weissman IL, et al. 2003. *Nature* 423: 448-52
114. Baeg GH, Selva EM, Goodman RM, Dasgupta R, Perrimon N. 2004. *Dev Biol* 276: 89-100
115. Takei Y, Ozawa Y, Sato M, Watanabe A, Tabata T. 2004. *Development* 131: 73-82
116. Tsuda M, Kamimura K, Nakato H, Archer M, Staatz W, et al. 1999. *Nature* 400: 276-80
117. Goto S, Taniguchi M, Muraoka M, Toyoda H, Sado Y, et al. 2001. *Nat Cell Biol* 3: 816-22
118. Hacker U, Lin X, Perrimon N. 1997. *Development* 124: 3565-73
119. Desbordes SC, Sanson B. 2003. *Development* 130: 6245-55
120. Lin X, Wei G, Shi Z, Dryer L, Esko JD, et al. 2000. *Dev Biol* 224: 299-311

121. Duncan G, McCormick C, Tufaro F. 2001. *J Clin Invest* 108: 511-6
122. Menon AK. 1995. *Trends Cell Biol* 5: 355-60
123. Wandall HH, Pedersen JW, Park C, Levery SB, Pizette S, et al. 2003. *J Biol Chem* 278: 1411-4
124. Schwientek T, Keck B, Levery SB, Jensen MA, Pedersen JW, et al. 2002. *J Biol Chem* 277: 32421-9
125. Haines, N., and Irvine, K. D. (2005) *Glycobiology* **15**, 335-346
126. Mucha J, Domlatil J, Lochnit G, Rendic D, Paschinger K, et al. 2004. *Biochem J* 382: 67-74
127. Kim BT, Tsuchida K, Lincecum J, Kitagawa H, Bernfield M, Sugahara K. 2003. *J Biol Chem* 278: 9116-24
128. Sugita M, Inagaki F, Naito H, Hori T. 1990. *J Biochem (Tokyo)* 107: 899-903
129. Sickmann T, Weske B, Dennis RD, Mohr C, Wiegandt H. 1992. *J Biochem (Tokyo)* 111: 662-9
130. Wandall, H. H., Pizette, S., Pedersen, J. W., Eichert, H., Levery, S. B., Mandel, U., Cohen, S. M., and Clausen, H. (2005) *J Biol Chem* **280**, 4858-4863
131. Schnaar RL, Collins BE, Wright LP, Kiso M, Tropak MB, et al. 1998. *Ann N Y Acad Sci* 845: 92-105
132. Perrimon N, Gans M. 1983. *Dev Biol* 100: 365-73
133. Perrimon N, Engstrom L, Mahowald AP. 1989. *Genetics* 121: 333-52
134. Shannon MP, Kaufman TC, Shen MW, Judd BH. 1972. *Genetics* 72: 615-38
135. Swan A, Hijal S, Hilfiker A, Suter B. 2001. *Genome Res* 11: 67-77
136. Rubsam R, Hollmann M, Simmerl E, Lammermann U, Schafer MA, et al. 1998. *Mech Dev* 72: 131-40
137. Goode S, Wright D, Mahowald AP. 1992. *Development* 116: 177-92
138. Goode S, Morgan M, Liang YP, Mahowald AP. 1996. *Dev Biol* 178: 35-50
139. Schupbach T. 1987. *Cell* 49: 699-707
140. Wieschaus E. 1978. *Results Probl Cell Differ* 9: 97-118
141. Harris RJ, Spellman MW. 1993. *Glycobiology* 3: 219-24
142. Sasamura T, Sasaki N, Miyashita F, Nakao S, Ishikawa HO, et al. 2003. *Development* 130: 4785-95
143. Shao L, Moloney DJ, Haltiwanger R. 2003. *J Biol Chem* 278: 7775-82
144. Lai EC. 2004. *Development* 131: 965-73
145. Irvine KD. 1999. *Curr Opin Genet Dev* 9: 434-41

146. Pourquie O. 1999. *Curr Opin Genet Dev* 9: 559-65
147. Wu JY, Rao Y. 1999. *Curr Opin Neurobiol* 9: 537-43
148. Roth S. 2001. *Curr Biol* 11: R779-81
149. Baron M, Aslam H, Flasza M, Fostier M, Higgs JE, et al. 2002. *Mol Membr Biol* 19: 27-38
150. Haltiwanger RS. 2002. *Curr Opin Struct Biol* 12: 593-8
151. de Celis JF, Garcia-Bellido A. 1994. *Mech Dev* 46: 109-22
152. de Celis JF, Garcia-Bellido A, Bray SJ. 1996. *Development* 122: 359-69
153. Diaz-Benjumea FJ, Cohen SM. 1993. *Cell* 75: 741-52
154. Milan M, Cohen SM. 2003. *Development* 130: 553-62
155. Panin VM, Papayannopoulos V, Wilson R, Irvine KD. 1997. *Nature* 387: 908-12
156. Okajima T, Irvine KD. 2002. *Cell* 111: 893-904
157. Haines N, Irvine KD. 2003. *Nat Rev Mol Cell Biol* 4: 786-97
158. Roth J, Wang Y, Eckhardt AE, Hill RL. 1994. *Proc Natl Acad Sci U S A* 91: 8935-9
159. Berger EG. 1999. *Biochim Biophys Acta* 1455: 255-68
160. Ten Hagen KG, Tran DT, Gerken TA, Stein DS, Zhang Z. 2003. *J Biol Chem* 278: 35039-48
161. Schwientek T, Bennett EP, Flores C, Thacker J, Hollmann M, et al. 2002. *J Biol Chem* 277: 22623-38
162. Nakamura N, Katano K, Toba S, Kurosaka A. 2004. *Biol Pharm Bull* 27: 1509-14
163. Guerardel Y, Balanzino L, Maes E, Leroy Y, Coddeville B, et al. 2001. *Biochem J* 357: 167-82
164. Bulet P, Dimarcq JL, Hetru C, Lagueux M, Charlet M, et al. 1993. *J Biol Chem* 268: 14893-7
165. Kramerov AA, Arbatsky NP, Rozovsky YM, Mikhaleva EA, Polesskaya OO, et al. 1996. *FEBS Lett* 378: 213-8
166. Ju T, Brewer K, D'Souza A, Cummings RD, Canfield WM. 2002. *J Biol Chem* 277: 178-86
167. Ju T, Cummings RD. 2002. *Proc Natl Acad Sci U S A* 99: 16613-8
168. Kudo T, Iwai T, Kubota T, Iwasaki H, Takayma Y, et al. 2002. *J Biol Chem* 277: 47724-31
169. Amerongen AV, Bolscher JG, Veerman EC. 1995. *Glycobiology* 5: 733-40
170. Jacob R, Alfalah M, Grunberg J, Obendorf M, Naim HY. 2000. *J Biol Chem* 275: 6566-72

171. Beckendorf SK, Kafatos FC. 1976. *Cell* 9: 365-73
172. Korayem AM, Fabbri M, Takahashi K, Scherfer C, Lindgren M, et al. 2004. *Insect Biochem Mol Biol* 34: 1297-304
173. Theopold U, Dorian C, Schmidt O. 2001. *Insect Biochem Mol Biol* 31: 189-97
174. Kramerova IA, Kramerov AA. 1999. *Dev Dyn* 216: 349-60
175. Hoffmann JA. 2003. *Nature* 426: 33-8
176. Bulet P, Hetru C, Dimarcq JL, Hoffmann D. 1999. *Dev Comp Immunol* 23: 329-44
177. McManus AM, Otvos L, Jr., Hoffmann R, Craik DJ. 1999. *Biochemistry* 38: 705-14
178. Ten Hagen KG, Tran DT. 2002. *J Biol Chem* 277: 22616-22
179. Flores C, Engels W. 1999. *Proc Natl Acad Sci U S A* 96: 2964-9
180. Topaz O, Shurman DL, Bergman R, Indelman M, Ratajczak P, et al. 2004. *Nat Genet* 36: 579-81
181. Toma V, Zuber C, Sata T, Roth J. 1999. *Glycobiology* 9: 1191-7
182. Hennet T, Hagen FK, Tabak LA, Marth JD. 1995. *Proc Natl Acad Sci U S A* 92: 12070-4
183. Xia L, Ju T, Westmuckett A, An G, Ivanciu L, et al. 2004. *J Cell Biol* 164: 451-9
184. Hennet T, Dinter A, Kuhnert P, Mattu TS, Rudd PM, Berger EG. 1998. *J Biol Chem* 273: 58-65
185. Perrimon N, Mohler D, Engstrom L, Mahowald AP. 1986. *Genetics* 113: 695-712
186. Yuan YP, Schultz J, Mlodzik M, Bork P. 1997. *Cell* 88: 9-11
187. Malissard M, Dinter A, Berger EG, Hennet T. 2002. *Eur J Biochem* 269: 233-9
188. Pearson WR. 1990. *Methods Enzymol* 183: 63-98
189. Brand AH, Perrimon N. 1993. *Development* 118: 401-15
190. Rong YS, Golic KG. 2001. *Genetics* 157: 1307-12
191. Kennerdell JR, Carthew RW. 1998. *Cell* 95: 1017-26
192. Lam G, Thummel CS. 2000. *Curr Biol* 10: 957-63
193. Wiegandt H. 1992. *Biochim Biophys Acta* 1123: 117-26
194. Giraudo CG, Maccioni HJ. 2003. *J Biol Chem* 278: 40262-71
195. Malinina L, Malakhova ML, Teplov A, Brown RE, Patel DJ. 2004. *Nature* 430: 1048-53
196. Marroquin LD, Elyassnia D, Griffitts JS, Feitelson JS, Aroian RV. 2000. *Genetics* 155: 1693-9
197. Khyami-Horani H. 2002. *J Basic Microbiol* 42: 105-10
198. Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, et al. 1998. *Microbiol Mol Biol Rev* 62: 775-806
199. Julenius, K., Molgaard, A., Gupta, R., and Brunak, S. (2005) *Glycobiology* **15**, 153-164

200. Hang HC, Yu C, Kato DL, Bertozzi CR. 2003. *Proc Natl Acad Sci U S A* 100: 14846-51
201. Perrimon N, Hacker U. 2004. *Development* 131: 2509-11; author reply 11-3
202. Rietveld A, Neutz S, Simons K, Eaton S. 1999. *J Biol Chem* 274: 12049-54
203. Simons K, Vaz WL. 2004. *Annu Rev Biophys Biomol Struct* 33: 269-95
204. Roepstorff K, Thomsen P, Sandvig K, van Deurs B. 2002. *J Biol Chem* 277: 18954-60
205. Arbeitman MN, Fleming AA, Siegal ML, Null BH, Baker BS. 2004. *Development* 131: 2007-21
206. Kohyama-Koganeya A, Sasamura T, Oshima E, Suzuki E, Nishihara S, et al. 2004. *J Biol Chem* 279: 35995-6002
207. Kawar ZS, Van Die I, Cummings RD. 2002. *J Biol Chem* 277: 34924-32
208. Lin X, Buff EM, Perrimon N, Michelson AM. 1999. *Development* 126: 3715-23
209. Luders F, Segawa H, Stein D, Selva EM, Perrimon N, et al. 2003. *Embo J* 22: 3635-44
210. Takemae H, Ueda R, Okubo R, Nakato H, Izumi S, et al. 2003. *J Biol Chem* 278: 15571-8
211. Kitagawa H, Egusa N, Tamura JI, Kusche-Gullberg M, Lindahl U, Sugahara K. 2001. *J Biol Chem* 276: 4834-8
212. The I, Perrimon N. 2000. *Nat Cell Biol* 2: E79-82
213. The I, Bellaiche Y, Perrimon N. 1999. *Mol Cell* 4: 633-9
214. Bellaiche Y, The I, Perrimon N. 1998. *Nature* 394: 85-8
215. Carroll SB, Winslow GM, Twombly VJ, Scott MP. 1987. *Development* 99: 327-32
216. Sen J, Goltz JS, Stevens L, Stein D. 1998. *Cell* 95: 471-81
217. Kamimura K, Fujise M, Villa F, Izumi S, Habuchi H, et al. 2001. *J Biol Chem* 276: 17014-21
218. Kamimura K, Rhodes JM, Ueda R, McNeely M, Shukla D, et al. 2004. *J Cell Biol* 166: 1069-79

PUBLICATIONS AND CONTRIBUTIONS

The following publications have appeared with data generated during this thesis:

[I] To the first publication: "*The Drosophila melanogaster brainiac protein is a glycolipid-specific beta 1,3N-acetylglucosaminyltransferase"*, I contributed all the results except those in Figure 2, "Structural analysis of BRN product" (by Fritz Altmann, Universität für Bodenkultur, Vienna).

[II] To the second publication: *"Resistance to a bacterial toxin is mediated by removal of a conserved glycosylation pathway required for toxin-host interactions"* I contributed Table 1, "Complementation of *D. melanogaster brn* $^{1.6P6}$ by *bre*-5".

[III] To the third publication : "*Characterization of mucin-type core-1 β1-3 galactosyltransferase homologous enzymes in Drosophila melanogaster*", I contributed all the results except Figure 2: Linkage analysis of the disaccharide Gal-GalNAc (by Fritz Altmann) and except Figure 4: HPLC profiling of glycolipid-derived oligosaccharides (by Andreas Hülsmeier, Institute of Physiology, University of Zurich).

Accelerated Publication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 36, Issue of September 6, pp. 32417–32420, 2002

The *Drosophila melanogaster brainiac* Protein Is a Glycolipid-specific β1,3*N*-Acetylglucosaminyltransferase*

Received for publication, June 26, 2002
Published, JBC Papers in Press, July 18, 2002,
DOI 10.1074/jbc.C200381200

Reto Müller, Friedrich Altmann‡, Dapeng Zhou§, and Thierry Hennet¶

From the Institute of Physiology, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland and the ‡Institute of Chemistry, Universität für Bodenkultur, Muthgasse 18, A-1190 Wien, Austria

Mutations at the *Drosophila melanogaster brainiac* locus lead to defective formation of the follicular epithelium during oogenesis and to neural hyperplasia. The *brainiac* gene encodes a type II transmembrane protein structurally similar to mammalian β1,3-glycosyltransferases. We have cloned the *brainiac* gene from *D. melanogaster* genomic DNA and expressed it as a FLAG-tagged recombinant protein in Sf9 insect cells. Glycosyltransferase assays showed that *brainiac* is capable of transferring *N*-acetylglucosamine (GlcNAc) to β-linked mannose (Man), with a marked preference for the disaccharide Man(β1,4)Glc, the core of arthro-series glycolipids. The activity of *brainiac* toward arthro-series glycolipids was confirmed by showing that the enzyme efficiently utilized glycolipids from insects as acceptors whereas it did not with glycolipids from mammalian cells. Methylation analysis of the *brainiac* reaction product revealed a β1,3 linkage between GlcNAc and Man, proving that *brainiac* is a β1,3GlcNAc-transferase. Human β1,3GlcNAc-transferases structurally related to *brainiac* were unable to transfer GlcNAc to Man(β1,4)Glc-based acceptor substrates and failed to rescue a homozygous lethal *brainiac* allele, indicating that these proteins are paralogous and not orthologous to *brainiac*.

The importance of glycoconjugates in regulating developmental processes is continually being supported by studies performed in various model organisms like *Caenorhabditis elegans* (1), *Drosophila melanogaster* (2), and the mouse (3). The *Drosophila* genes *sugarless*, *sulfateless*, *pipe, tout-velu*, and *dally* participate in the formation of proteoglycans. Loss of function mutations in some of these genes produce polarity phenotypes mechanistically connected to incorrect diffusion of the signaling proteins *wingless* and *hedgehog* (4–6). The *rotated abdomen* locus, whose disruption is associated with a helical rotation of the body, has been found to encode a potential *O*-mannosyltransferase (7), and *fringe*, which modulates the interaction of the *Notch* receptor with its ligands (8), has recently been demonstrated to be a β1,3*N*-acetylglucosaminyltransferase (GlcNAcT)[1] (9, 10).

The *Drosophila* gene *brainiac* (*brn*) (11) encodes a protein that shares structural motifs with β1,3glycosyltransferases (12, 13). The *brn* gene is localized on the X chromosome. *brn* was shown to cooperate with the epidermal growth factor receptor and one of its ligands, the *Drosophila* TGFα homologue *gurken* (11) during oogenesis. Mutant *brn* alleles exhibit altered morphology of the follicular epithelium (11), female sterility (14), and germ line loss (15). Furthermore, *brn* embryos develop neural hyperplasia and epidermal hypoplasia (11) as encountered with *Notch* hypomorphic alleles and other neurogenic mutants, suggesting implications of *brn* in *Notch* signaling (16, 17).

While the relationships between *brn* and specific signaling pathways have been examined genetically, the nature of these interactions remained elusive as long as the biochemical function of *brn* was unclear. In the present study, we show that *brn* has a β1,3*N*-acetylglucosaminyltransferase (GlcNAcT) activity directed toward the Man(β1,4)Glc core structure of arthro-series glycolipids.

EXPERIMENTAL PROCEDURES

Cloning and Expression—The *brn* gene was amplified by PCR from *D. melanogaster OregonR* genomic DNA during 30 cycles at 95 °C for 45 s, 55 °C for 30 s, 72 °C for 60 s using the primers 5′-TTTGGATCCGTCGCCATGCAAAGT-3′ and 5′-CCTGTTCTAGATGCTACGCGTAAT-3′. The resulting 1.0-kb fragment was digested with *Bam*HI and *Xba*I and subcloned into the pFastbac-FLAG(a) vector (Invitrogen) linearized at the *Bam*HI and *Xba*I sites. The FLAG-tagged *brn* gene was expressed as a recombinant baculovirus in insect cells as described previously (13). Infected cells (10^7) were lysed at 72 h post-infection in 600 μl of 50 mM Tris/HCl, pH 7.4, 150 mM NaCl, 1% Triton X-100, and a protein inhibitor mixture (*complete*, EDTA free, Roche Diagnostics) on ice. Post-nuclear supernatants were incubated with 240 μl of anti-FLAG M2-agarose beads (Sigma) under rotation for 2.5 h at 4 °C. Beads were washed three times with Tris-buffered saline and used as enzyme source for assays.

Glycosyltransferase Assays—All donor and acceptor substrates were from Sigma except Man(β1,4)Glc(β1-OpNP) (pNP = *p*-nitrophenyl), which was purchased from Toronto Research (North York, Canada). Glycosyltransferase activity was assayed for 60 min at 25 °C with 15 μl of beads, 5% Me_2SO, 20 mM $MnCl_2$, 0.08 mM UDP-GlcNAc including 5×10^4 cpm of UDP-[^{14}C]GlcNAc (Amersham Biosciences), and various acceptors (see Table I). Reaction products were purified over C_{18} Sep-Pak cartridges (Waters) (18) and quantified in a Tri-Carb 2900TR liquid scintillation counter (Packard) with luminescence correction.

Glycolipid Extraction—*D. melanogaster* Schneider 2 cells, *Spodoptera frugiperda* Sf9 cells, and human colon carcinoma Caco-2 cells were washed three times in phosphate-buffered saline and extracted in isopropanol:hexane:H_2O (55:25:20). Extracts were spun twice at 500 *x g*, and supernatants were dried under N_2. Phospholipids were removed by saponification in 0.2 M NaOH in methanol for 24 h at 37 °C. After neutralization with HCl, the extracts were expanded to theoretical upper phase (methanol:water:chloroform, 47:48:3), applied on C_{18} Sep-

* This work was supported by Swiss National Science Foundation Grant 631-062662.00 (to T. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: Dept. of Molecular Biology, Princeton University, Washington Rd., Princeton, NJ 08544-1014.

¶ To whom correspondence should be addressed: Inst. of Physiology, Winterthurerstrasse 190, 8057 Zurich, Switzerland. Tel.: 41-1-635-5080; Fax: 41-1-635-6814; E-mail: thennet@access.unizh.ch.

[1] The abbreviations used are: GlcNAcT, *N*-acetylglucosaminyltransferase; *brn*, *brainiac*; TLC, thin-layer chromatography; pNP, *p*-nitrophenyl; Cer, ceramide; Lc2, Gal(β1,4)Glc-Cer; Lc3, GlcNAc(β1,3)Gal(β1,4)Glc-Cer; HPLC, high performance liquid chromatography.

Pak cartridges, and eluted with 5 ml of methanol. Eluates were dried under N_2 and resuspended in 500 μl of methanol. The procedure yielded about 120 μg of mannose equivalents for 10^8 S2 and Sf9 cells and 20 μg of mannose equivalents from 10^7 Caco-2 cells as determined by the phenol sulfuric acid assay (19).

Thin-layer Chromatography (TLC)—Glycolipids (5 μg of mannose equivalents per assay) were dried under N_2 and incubated together with 10 μl of beads-bound enzyme in 50 μl of 50 mM cacodylate buffer, pH 7.1, 20 mM $MnCl_2$, 0.06% Triton X-100, 2.5×10^4 cpm of UDP-[^{14}C]GlcNAc for 90 min at 25 °C. Reaction products were expanded to theoretical upper phase and purified over C_{18} Sep-Pak cartridges as described above. After drying over N_2, the eluates were taken up in 100 μl of methanol:chloroform (1:1) and separated on aluminum high-performance thin-layer chromatography plates (Merck, Darmstadt, Germany) using a solvent system of chloroform:methanol:0.25% $CaCl_2$ (5:4:1). Plates were stained with orcinol sulfuric acid (Sigma). The [^{14}C]GlcNAcβ1,3Galβ1,4Glc-ceramide (Lc3) standard was produced enzymatically with the Lc3 synthase β1,3 GlcNacT protein (20) using Galβ1,4Glc-ceramide (Lc2) (Sigma) as acceptor substrate.

brn Complementation in Drosophila—Human β3GnT1 (21), β3GnT4 (22) and β3GnT5 (20) cDNAs and the *Drosophila brn* gene were subcloned into the pUAST vector (23). The rescue constructs pUAST-β3GnT1, pUAST-β3GnT4, pUAST-β3GnT5, and pUAST-*brn* were injected together with the pUChspΔ2–3 P-element helper plasmid (Flybase accession FBmc0000938) into *yellow white Drosophila* embryos using standard procedures. Then, *white*$^+$ progeny was selected and X-chromosomal insertions of the transgene excluded. The GAL4 lines, driving ubiquitous expression of the UAS-transgenes in an *armadillo* pattern (24), carry Bloomington Stock numbers 1560 and 1561. Males of the genotype *yellow white*/Y; transgene/+ were mated to virgins *forked brn*$^{1.6P6}$/FM6-w^1; 1560 GAL4/+ and *forked brn*$^{1.6P6}$/FM6-w^1; 1561 GAL4/+ and the progeny examined for males carrying the *forked* mutation for 8 days after eclosion of the first flies. At least two independent lines of each transgene were used for the complementation assay, which were repeated four times.

Structural Analysis—A mixture of substrate and of 10 nmol of product was separated by reversed phase HPLC on a 3 × 250 mm column filled with 5 μM ODS Hypersil (Shandon) at a flow rate of 0.6 ml/min. The column was eluted with a linear gradient from 6 to 24% of methanol during 18 min in 0.1 M ammonium acetate, pH 4.0. *p*-Nitrophenylglycosides were monitored at 245 nm. The mixture was also analyzed after incubation with *N*-acetyl-β-hexosaminidase from jack beans (Sigma) (25). The fraction of interest was collected in a screw capped glass vial and dried in a speed-vac concentrator. A small aliquot was used for matrix assisted laser desorption mass spectrometry as described elsewhere (25). The sample was dried over phosphorus pentoxide *in vacuo* and permethylated using NaOH (26). Partially permethylated alditol acetates were prepared using $NaBD_4$ as the reducing agent and analyzed by gas chromatography/mass spectrometry using a 60 m SP2330 (Restek) (27) and a Finnigan Ion Trap ITD800. Derivatives of terminal and 3-substituted galactose served to compare retention times with the data given by Doares *et al.* (27).

RESULTS

We have cloned the *D. melanogaster brn* gene by PCR amplification and expressed it as an N-terminally FLAG-tagged full-length protein in Sf9 insect cells. The recombinant *brn* protein was bound to anti-FLAG-agarose beads, and cellular contaminants such as possible endogenous acceptor substrates were washed out before assaying for enzymatic activity. A GlcNAcT activity was only detected toward the Manβ1-OpNP) acceptor when monosaccharide substrates were assayed (Table I). Highest activity was measured toward the disaccharide acceptor Manβ1,4Glcβ1-OpNP), whereas a slight activity was also detected toward Galβ1,4Glcβ1-OpNP) (Table I). The Manβ1,4Glc structure represents the core of arthro-series glycolipids found in nematodes (28) and insects (29) among others.

In *Drosophila*, the arthro-series Manβ1,4Glc core is elongated with a β1,3-linked GlcNAc (30), suggesting that *brn* may represent the enzyme catalyzing this step. To test this hypothesis, we have isolated neutral glycolipids from *Drosophila* S2 and *Spodoptera* Sf9 cells and assayed these glycolipids as acceptors for the anti-FLAG beads-bound *brn* enzyme. A significant GlcNAc-transferase activity was detected when incubating *brn* together with insect glycolipids, whereas only a low activity was measured with glycolipids extracted from mammalian Caco-2 cells, likely reflecting the low specificity of *brn* for lactosylceramide. The reaction products were separated by TLC and plates were autoradiographed, revealing a [^{14}C]GlcNAc-labeled band at the size of a trihexoside ceramide in S2 and Sf9 cells (Fig. 1).

TABLE I
Acceptor substrate specificity of brn

Acceptor substrate (20 mM)	Mock[a]	*brn*[b]
	pmol/min/ml	
Glcα1-OpNP)	7.2	5.2
Glcβ1-OpNP)	5.4	5.7
Galα1-OpNP)	8.4	11.1
Galβ1-OpNP)	6.2	12.6
GalNAcα1-OpNP)	6.7	7.2
GalNAcβ1-OpNP)	5.3	8.9
Fucα1-OpNP)	7.6	4.3
Fucβ1-OpNP)	6.6	5.5
Manα1-OpNP)	4.6	16.1
Manβ1-OpNP)	5.6	400.0
Manβ1,4Glcβ1-OpNP)	3.0	855.1
Galβ1,4Glcβ1-OpNP)	4.5	24.9

[a] Anti-FLAG bead bound lysate from Sf9 cells infected with mock baculovirus.
[b] Anti-FLAG bead bound lysate from Sf9 cells infected with *brn* baculovirus.

The nature of the linkage between GlcNAc and the underlying β-linked Man residue was investigated by methylation analysis of the *brn* reaction product GlcNAc-Manβ1-OpNP). In reversed phase HPLC, the presumed disaccharide product eluted slightly ahead of the substrate Manβ1-OpNP). The disaccharide peak disappeared upon incubation with *N*-acetyl-β-hexosaminidase (Fig. 2*A*). The purified fraction corresponding to the disaccharide peak exhibited a pseudomolecular ion of *m/z* 513.5. Linkage analysis of the GlcNAc-Manβ1-OpNP) disaccharide product gave a peak at the relative retention time of 0.597, which suggests a 2- or a 3-substituted mannosyl residue (27). The fragment spectrum clearly identified the derivative as substituted in the 3-position (Fig. 2*B*), thus confirming the identity of *brn* as a β1,3 GlcNAcT.

The *brn* protein is structurally related to human β1,3 glycosyltransferase enzymes (12, 13). The acceptor specificity of *brn* for the arthro-series glycolipid core suggested that it represents a paralogous enzyme to the mammalian β1,3 glycosyltransferases, including β1,3 galactosyltransferases (13, 31, 32), β1,3 GlcNAcT (20, 22), and a β1,3-*N*-acetylgalactosaminyltransferase (33) acting on GlcNAcβ-, Galβ-, and GalNAcβ-based acceptors. Although no mammalian β1,3 GlcNAcT has been described to act on β-linked Man acceptors, we have investigated whether the three human β1,3 GlcNAcT structurally closest to *brn* can complement the lethal phenotype of *brn* deficient *Drosophila* flies. To this end, we have expressed the human β3GnT-I (21), -IV (22), and -V (20) in *brn*$^{1.6P6}$ mutant flies (34) using the UAS-GAL4 transgenesis system (23).

The human β3GnT transgenes and a *brn* transgene were expressed in flies carrying the allele *brn*$^{1.6P6}$, which causes lethality at the late pupal stage. The transgenes were expressed ubiquitously using *armadillo* GAL4 transactivator lines. The *brn* transgene did rescue *brn*$^{1.6P6}$ mutant males from their hemizygous late pupal lethality, whereas the human β3GnT transgenes did not (Table II). The rescue of *brn*$^{1.6P6}$ males was confirmed by detection of the *forked* marker, whose gene is located besides the *brn*$^{1.6P6}$ allele on the X chromosome. Control crosses of females carrying *brn*$^{1.6P6}$ with *yellow white* males did not yield any living *brn*$^{1.6P6}$ *forked*/Y males either. The inability of human β3GnT enzymes to compensate for the

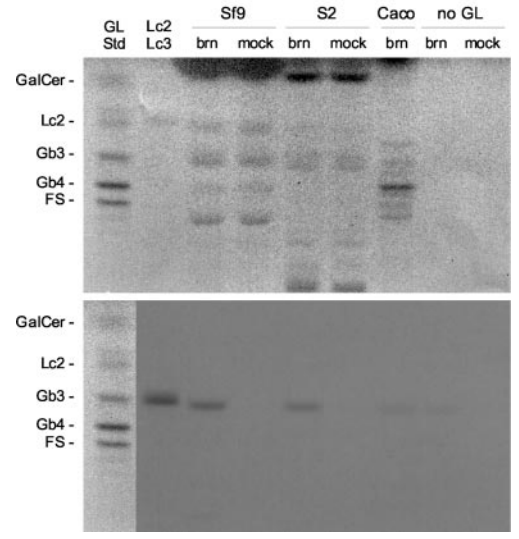

FIG. 1. **TLC separation of *brn*-modified glycolipids.** The *top panel* shows a plate stained with orcinol reagent and the *bottom panel* the autoradiogram of the same plate. *brn*-bound beads (*brn*) or beads preincubated with mock-infected Sf9 cells (*mock*) were incubated with neutral glycolipids from *Drosophila* S2 cells (S2), *Spodoptera* Sf9 cells (*Sf9*), human Caco-2 cells (*Caco*), or without added glycolipids (*no GL*). The neutral glycolipid standard (*GL Std*) contained: Gal-Cer; Gal(β1,4)Glc-Cer (*Lc2*); Gal(α1,4)Gal(β1,4)Glc-Cer (*Gb3*); Gal(α1,4)Gal(β1,4)Glc-Cer (*Gb4*), and GalNAc(α1,3)GalNAc(β1,3)Gal(α1,4)Glc-Cer (*FS*). *Lc2/Lc3*, Lc2 was elongated to Lc3 by incorporation of [^{14}C]GlcNAc catalyzed by the human β3GnT-V enzyme (20).

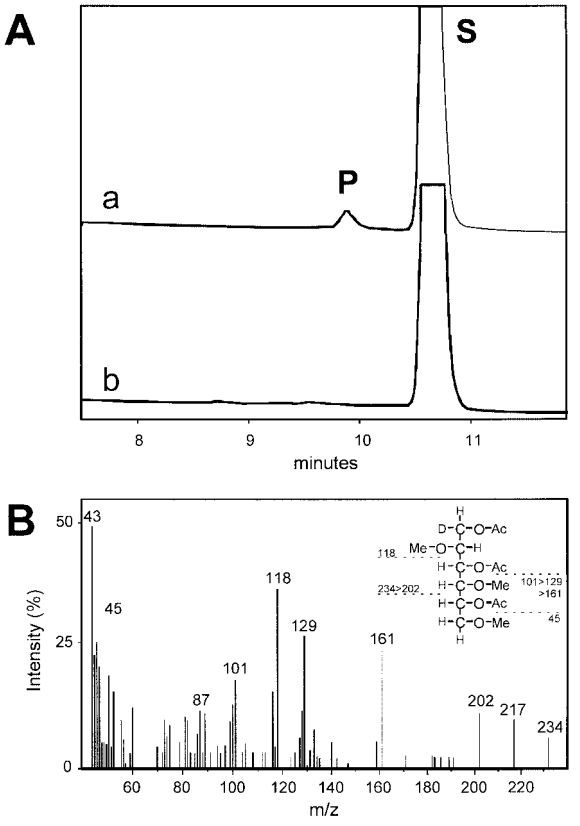

FIG. 2. **Structural analysis of *brn* product.** *A*, HPLC separation of product and substrate of β1,3GlcNAcT *brn*. *p*-Nitrophenyl-β-D-mannopyranoside was incubated with *brn* in the presence of UDP-GlcNAc. The mixture was prepurified over Sep-Pak C_{18} cartridges and subjected to reversed phase chromatography (*trace a*). A small product peak (*P*) eluted ahead of the substrate (*S*). The product disappeared upon incubation with β-*N*-acetylhexosaminidase (*trace b*) which indicates it to contain a β-linked GlcNAc residue. *B*, methylation analysis of *p*-nitrophenyl disaccharide. The electron impact mass spectrum of the partially methylated monodeuterated alditol acetate derived from the mannosyl residue of the disaccharide product shows several fragments indicative of a substitution in position 3 as depicted by the insert. Especially the presence of mass 118 and the absence of mass 190 exclude a 2-substitution, which could not be ruled out from the retention time alone (27).

loss of *brn* activity in mutant flies suggested that the former enzymes cannot elongate the arthro-series glycolipid core *in vivo*. This was confirmed *in vitro* by showing that the human β3GnT enzymes did not exhibit significant activity toward the Man(β1,4)Glc(β1-OpNP) acceptor (Table II).

DISCUSSION

We have shown that *Drosophila brn*, a member of the β1,3 glycosyltransferase family, encodes a β1,3 GlcNAcT enzyme with a specificity for the Man(β1,4)Glc disaccharide found in arthro-series glycolipids (29). Several mammalian enzymes structurally related to *brn* have been suggested to represent homologues (35–37). However, the specificity of *brn* for Man(β1,4)Glc, a disaccharide that has never been described in vertebrates, rather indicates that *brn* and mammalian β1,3 glycosyltransferases are paralogous proteins derived from a common ancestor gene.

The functional disparity between the β1,3 GlcNAcT *brn* and mammalian β1,3 GlcNAcT enzymes is further supported by the inability of the latter to complement the lethal phenotype of the mutant allele *brn$^{1.6P6}$* in *Drosophila*. The specificity of *brn* toward Manβ1,4Glc-Cer suggests the presence of functional homologues only in organisms harboring arthro-series glycolipids, whose core structure is GlcNAc(β1,3)Man(β1,4)Glc-Cer. A protein structurally related to *brn* has recently been described in *C. elegans* (38), which express arthro-series glycolipids (28). The loss of that gene, named *bre-5* (39), renders the animal resistant to high doses of *Bacillus thuringiensis* Bt toxin. Since Bt toxin binds to arthro-series glycolipids (40), it is possible that *bre-5* participates in the formation of this class of glycolipids in *C. elegans* and thereby represents a true orthologue of *brn*.

brn mutations affect follicle cell-germ line interactions and lead to neurogenic phenotypes in *Drosophila* embryos. Considering the involvement of *brn* in glycolipid biosynthesis, one can envision that arthro-series glycolipids may regulate cell adhesion, proliferation, and differentiation via carbohydrate-lectin interactions. On the other hand, arthro-series glycolipids may modulate specific signaling proteins in a way similar to gangliosides affecting the epidermal growth factor receptor (41, 42), insulin receptor (43), and platelet-derived growth factor receptor (44) signaling cascades. The notion that *brn* glycolipid products interact with adhesion or signaling proteins implies

TABLE II
Complementation of Drosophila brn$^{1.6P6}$

Rescue of the *brn$^{1.6P6}$* late pupal lethal phenotype by ubiquitous expression of *Drosophila brn* and human β1,3GlcNAcT transgenes.

β1,3GlcNAcT gene[a]	Lines[b]	*brn$^{1.6P6}$* rescue[c]	GlcNAc->Man(β1,4)Glc activity[d]
			%
Drosophila brn	2	2	100
Human β3GnT-I	2	0	6.3
Human β3GnT-IV	2	0	1.6
Human β3GnT-V	6	0	8.6

[a] β3GnT-I (21), β3GnT-IV (22), and β3GnT-V (20).
[b] Number of independent lines per transgene tested.
[c] Number of independent lines per transgene rescueing *brn$^{1.6P6}$* (34). Rescue was scored by counting males alive with a *forked* phenotype within 8 days after eclosion of control animals.
[d] *In vitro* GlcNAc-transferase activity towards Man(β1,4)Glc(β1-OpNP) given in percentage of the activity measured with *brn*.

that other mutant genes with phenotypes similar to those encountered in *brn* mutant flies may encode partner lectin/signaling proteins. Along this line, *Drosophila egghead* mutants have similar and non-additive phenotypes to *brn* (17). Experiments aimed at characterizing the biochemical and functional relation between *brn* products and the *egghead* protein should reveal the mechanisms how arthro-series glycolipids regulate morphogenic events during *Drosophila* development.

Acknowledgments—We thank Erich Frei, Michael Daube, and Markus Noll from the Institute of Molecular Biology at the University of Zürich for their assistance with the transgenic expression in *Drosophila*.

REFERENCES

1. Herman, T., and Horvitz, H. R. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 974–979
2. Perrimon, N., and Bernfield, M. (2000) *Nature* **404,** 725–728
3. Hennet, T., and Ellies, L. G. (1999) *Biochim. Biophys. Acta* **1473,** 123–136
4. Sen, J., Goltz, J. S., Stevens, L., and Stein, D. (1998) *Cell* **95,** 471–481
5. Hacker, U., Lin, X., and Perrimon, N. (1997) *Development* (*Camb.*)**124,** 3565–3573
6. Bellaiche, Y., The, I., and Perrimon, N. (1998) *Nature* **394,** 85–88
7. Martin-Blanco, E., and Garcia-Bellido, A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 6048–6052
8. Panin, V. M., Papayannopoulos, V., Wilson, R., and Irvine, K. D. (1997) *Nature* **387,** 908–912
9. Moloney, D. J., Panin, V. M., Johnston, S. H., Chen, J. H., Shao, L., Wilson, R., Wang, Y., Stanley, P., Irvine, K. D., Haltiwanger, R. S., and Vogt, T. F. (2000) *Nature* **406,** 369–375
10. Bruckner, K., Perez, L., Clausen, H., and Cohen, S. (2000) *Nature* **406,** 411–415
11. Goode, S., Wright, D., and Mahowald, A. P. (1992) *Development* (*Camb.*) **116,** 177–192
12. Yuan, Y. P., Schultz, J., Mlodzik, M., and Bork, P. (1997) *Cell* **88,** 9–11
13. Hennet, T., Dinter, A., Kuhnert, P., Mattu, T. S., Rudd, P. M., and Berger, E. G. (1998) *J. Biol. Chem.* **273,** 58–65
14. Gans, M., Audit, C., and Masson, M. (1975) *Genetics* **81,** 683–704
15. Swan, A., Hijal, S., Hilfiker, A., and Suter, B. (2001) *Genome Res.* **11,** 67–77
16. Goode, S., and Perrimon, N. (1997) *Cold Spring Harbor Symp. Quant. Biol.* **62,** 177–184
17. Goode, S., Melnick, M., Chou, T. B., and Perrimon, N. (1996) *Development* (*Camb.*) **122,** 3863–3879
18. Malissard, M., Borsig, L., Di Marco, S., Grutter, M. G., Kragl, U., Wandrey, C., and Berger, E. G. (1996) *Eur. J. Biochem.* **239,** 340–348
19. Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F. (1956) *Anal. Chem.* **28,** 350–356
20. Henion, T. R., Zhou, D., Wolfer, D. P., Jungalwala, F. B., and Hennet, T. (2001) *J. Biol. Chem.* **276,** 30261–30269
21. Zhou, D., Dinter, A., Gutierrez Gallego, R., Kamerling, J. P., Vliegenthart, J. F. G., Berger, E. G., and Hennet, T. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 406–411, correction (2000) Proc. Natl. Acad. Sci. U. S. A. **97,** 11673–11675
22. Shiraishi, N., Natsume, A., Togayachi, A., Endo, T., Akashima, T., Yamada, Y., Imai, N., Nakagawa, S., Koizumi, S., Sekine, S., Narimatsu, H., and Sasaki, K. (2001) *J. Biol. Chem.* **276,** 3498–3507
23. Brand, A. H., and Perrimon, N. (1993) *Development* (*Camb.*) **118,** 401–415
24. Sanson, B., White, P., and Vincent, J. P. (1996) *Nature* **383,** 627–630
25. Wilson, I. B., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C. J., Kamerling, J. P., and Altmann, F. (2001) *Glycobiology* **11,** 261–274
26. Ciucanu, I., and Kerek, F. (1984) *Carbohydr. Res.* **131,** 209–217
27. Doares, S. H., Albersheim, P., and Darvill, A. G. (1991) *Carbohydr. Res.* **210,** 311–317
28. Gerdt, S., Lochnit, G., Dennis, R. D., and Geyer, R. (1997) *Glycobiology* **7,** 265–275
29. Wiegandt, H. (1992) *Biochim. Biophys. Acta* **1123,** 117–126
30. Seppo, A., Moreland, M., Schweingruber, H., and Tiemeyer, M. (2000) *Eur. J. Biochem.* **267,** 3549–3558
31. Miyazaki, H., Fukumoto, S., Okada, M., Hasegawa, T., and Furukawa, K. (1997) *J. Biol. Chem.* **272,** 24794–24799
32. Kolbinger, F., Streiff, M. B., and Katopodis, A. G. (1998) *J. Biol. Chem.* **273,** 433–440
33. Okajima, T., Nakamura, Y., Uchikawa, M., Haslam, D. B., Numata, S. I., Furukawa, K., Urano, T., and Furukawa, K. (2000) *J. Biol. Chem.* **275,** 40498–40503
34. Perrimon, N., Engstrom, L., and Mahowald, A. P. (1989) *Genetics* **121,** 333–352
35. Egan, S., Cohen, B., Sarkar, M., Ying, Y., Cohen, S., Singh, N., Wang, W., Flock, G., Goh, T., and Schachter, H. (2000) *Glycoconj. J.* **17,** 867–875
36. Cole, S. E., Mao, M. S., Johnston, S. H., and Vogt, T. F. (2001) *Mamm. Genome* **12,** 177–179
37. Vollrath, B., Fitzgerald, K. J., and Leder, P. (2001) *Mol. Cell. Biol.* **21,** 5688–5697
38. Griffitts, J. S., Whitacre, J. L., Stevens, D. E., and Aroian, R. V. (2001) *Science* **293,** 860–864
39. Marroquin, L. D., Elyassnia, D., Griffitts, J. S., Feitelson, J. S., and Aroian, R. V. (2000) *Genetics* **155,** 1693–1699
40. Dennis, R. D., Wiegandt, H., Haustein, D., Knowles, B. H., and Ellar, D. J. (1986) *Biomed. Chromatogr.* **1,** 31–37
41. Bremer, E. G., Schlessinger, J., and Hakomori, S. (1986) *J. Biol. Chem.* **261,** 2434–2440
42. Miljan, E. A., Meuillet, E. J., Mania-Farnell, B., George, D., Yamamoto, H., Simon, H. G., and Bremer, E. G. (2002) *J. Biol. Chem.* **277,** 10108–10113
43. Nojiri, H., Stroud, M., and Hakomori, S. (1991) *J. Biol. Chem.* **266,** 4531–4537
44. Van Brocklyn, J., Bremer, E. G., and Yates, A. J. (1993) *J. Neurochem.* **61,** 371–374

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 46, Issue of November 14, pp. 45594–45602, 2003
Printed in U.S.A.

Resistance to a Bacterial Toxin Is Mediated by Removal of a Conserved Glycosylation Pathway Required for Toxin-Host Interactions*

Received for publication, July 25, 2003, and in revised form, August 22, 2003
Published, JBC Papers in Press, August 27, 2003, DOI 10.1074/jbc.M308142200

Joel S. Griffitts‡, Danielle L. Huffman‡, Johanna L. Whitacre‡, Brad D. Barrows‡, Lisa D. Marroquin‡, Reto Müller§, Jillian R. Brown¶, Thierry Hennet§, Jeffrey D. Esko¶, and Raffi V. Aroian‡‖

From the ‡Section of Cell and Developmental Biology, University of California, San Diego, La Jolla, California 92093, §Institute of Physiology, University of Zürich, Winterhurerstrasse 190, 8057 Zürich, Switzerland, and ¶Department of Cellular and Molecular Medicine, Glycobiology Research and Training Center, University of California, San Diego, La Jolla, California 92093

Crystal (Cry) proteins made by the bacterium *Bacillus thuringiensis* are pore-forming toxins that specifically target insects and nematodes and are used around the world to kill insect pests. To better understand how pore-forming toxins interact with their host, we have screened for *Caenorhabditis elegans* mutants that resist Cry protein intoxication. We find that Cry toxin resistance involves the loss of two glycosyltransferase genes, *bre-2* and *bre-4*. These glycosyltransferases function in the intestine to confer susceptibility to toxin. Furthermore, they are required for the interaction of active toxin with intestinal cells, suggesting they make an oligosaccharide receptor for toxin. Similarly, the *bre-3* resistance gene is also required for toxin interaction with intestinal cells. Cloning of the *bre-3* gene indicates it is the *C. elegans* homologue of the *Drosophila egghead* (*egh*) gene. This identification is striking given that the previously identified *bre-5* has homology to *Drosophila brainiac* (*brn*) and that *egh-brn* likely function as consecutive glycosyltransferases in *Drosophila* epithelial cells. We find that, like in *Drosophila*, *bre-3* and *bre-5* act in a single pathway in *C. elegans*. *bre-2* and *bre-4* are also part of this pathway, thereby extending it. Consistent with its homology to *brn*, we demonstrate that *C. elegans bre-5* rescues the *Drosophila brn* mutant and that BRE-5 encodes the dominant UDP-GlcNAc:Man GlcNAc transferase activity in *C. elegans*. Resistance to Cry toxins has uncovered a four component glycosylation pathway that is functionally conserved between nematodes and insects and that provides the basis of the dominant mechanism of resistance in *C. elegans*.

Bacterial pore-forming toxins that damage membranes are important virulence factors associated with pathogenic bacteria (1, 2). Examples include aerolysin from *Aeromonas hydrophila*, α toxin from *Staphylococcus aureus*, hemolysin from *Escherichia coli*, and main family crystal (Cry)[1] proteins from *Bacillus thuringiensis*. Of these, Cry proteins are unique in that they specifically target invertebrates (insects and nematodes) and are generally considered innocuous to mammals. Mammalian resistance to Cry toxins is thought to be due to toxin insolubility, lack of proper proteolytic processing, and lack of proper receptors in the mammalian gut (3). As such, Cry proteins demonstrate a degree of safety toward vertebrates that is unmatched by any other pest control product and are used extensively around the world in the control of insect pests that damage crops and carry disease.

We previously reported on the isolation of *Caenorhabditis elegans bre* mutants (for *Bt*-toxin resistant) that resist *Bt* Cry protein intoxication (4). When fed the main family crystal protein Cry5B either as protein crystals or as *E. coli* produced protein, *C. elegans* becomes rapidly sick and shows degeneration of the intestine, loss of coloration, inhibition of growth, inhibition of progeny production, and eventual death (4, 5). Cry5B-resistant *bre* mutants were isolated among the F2 progeny of ethylmethanesulfonate mutagenized larvae. After screening 42,000 mutagenized haploid genomes for animals resistant to Cry5B, we isolated 14 alleles of *bre-2*, 19 alleles of *bre-3*, 9 alleles of *bre-4*, and 2 alleles of *bre-5* (4).[2] All are strongly resistant to the pore-forming toxin, and all are healthy in the absence of toxin as well. To date we have also isolated one allele of *bre-1* that has significantly weaker resistance to Cry5B (4). No other loci have been identified in these screens. Thus, strong, healthy resistance to the pore-forming toxin inevitably leads to mutation of *bre-2*, *bre-3*, *bre-4*, or *bre-5*. The identification of *bre-5* as a β1,3-glycosyltransferase gene has been reported (6).

To our knowledge this approach is unique in its objective to genetically characterize how whole animals interact with and resist a bacterial pore-forming toxin. In the case of Cry proteins, identifying the genes that mutate to confer resistance is important since the major threat to the long term utility of insecticidal Cry proteins is the emergence of resistance. More generally, understanding the pathways that can mediate resistance might suggest therapeutic strategies for coping with bacterial toxins that target mammals.

Here, we identify the Cry5B toxin resistance genes *bre-2* and *bre-4* as glycosyltransferase genes that previously have not

* This work was funded by National Science Foundation Grant MCB-9983013 and grants from the Burroughs-Wellcome Foundation and the Beckman Foundation (to R. V. A.) and by National Institutes of Health Grant GM33063 (to J. D. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed: Section of Cell and Developmental Biology, 9500 Gilman Dr., Dept. 0349, Bonner Hall 4430, La Jolla, CA 92093-0349. Tel.: 858-822-1396; Fax: 858-822-0808; E-mail: raroian@ucsd.edu.

[1] The abbreviations used are: Cry, crystal; ORF, open reading frame; HA, hemagglutinin.

[2] L. D. Marroquin and R. V. Aroian, unpublished results.

been genetically characterized. We find that *bre-2* and *bre-4* function in the intestine to confer susceptibility to Cry5B and are also required for the interaction of toxin with intestinal cells in *vivo*. We also identify the *bre-3* resistance gene as sharing significant homology to *Drosophila egghead* (*egh*). Because *bre-5* shares sequence similarity to the *Drosophila brainiac* (*brn*) gene and since *egh* and *brn* function in a single pathway in *Drosophila*, we hypothesized that the Cry protein intoxication pathway in *C. elegans* is equivalent to the *Drosophila* pathway. We indeed show that *bre-3* and *bre-5* function in a common pathway in *C. elegans*, that *bre-2* and *bre-4* are also part of this pathway, and that the *C. elegans bre-5* gene has the same in *vivo* and in *vitro* activities as *Drosophila brn*.

EXPERIMENTAL PROCEDURES

C. elegans Culture and Microscopy—*C. elegans* was propagated using standard techniques with strain Bristol N2 as the wild type (7). All nematode assays were carried out at 20 °C. For microscopy of live animals, nematodes were mounted on pads consisting of 2% agarose and 0.1% sodium azide. Low magnification microscopy was carried out on an Olympus BX-60 microscope using a 10× objective coupled to a 0.5× camera mount and a DVC camera. Endocytosis assays were imaged and deconvolved using a DeltaVision system (Applied Precision) on an Olympus IX-70 microscope using a 40× objective (NA1.35) and differential interference contrast or fluorescence optics. Immunofluorescence images were captured on the DeltaVision system using a 40× or 100× objective (NA1.35) and differential interference contrast or fluorescence optics.

Cloning of bre Genes—*bre-2*(*ye31*) was placed in *trans* to *dpy-18*(*e364*) *unc-64*(*e246*). 25/25 Dpy nonUncs segregated *bre-2*(*ye31*), and 23/23 Unc nonDpys segregated *bre*(+). The deficiency *eDf2* also failed to complement *bre-2*(*ye31*), suggesting that *bre-2* resided between *unc-64* and the right breakpoint of *eDf2*. The *bre-2*(*ye71*) allele was generated by mutagenesis in the Hawaiian strain background (CB4856) and then placed in *trans* to *unc-25*(*e156*) *bli-5*(*e518*). Unc nonDpy and Dpy non-Unc recombinants were used to follow the segregation of single nucleotide polymorphisms, which indicated that *bre-2* mapped between nucleotides 3,645 and 153,209 of Y39E4B. This region contains 14 predicted open reading frames (ORFs). Sequencing of cDNAs isolated from two *bre-2* alleles indicated that both alleles contained point mutations in Y39E4B.9. Identification of the putative glycosyltransferase Y39E4B.9 as *bre-2* was confirmed by reproducing the phenotype using RNA-mediated interference (data not shown) and by cDNA rescue from a *bre-2* mutant. Analogous to *bre-3* (below, including spliced leader 1 primer at the 5′ end), a complete *bre-2* cDNA was constructed using reverse-transcribed RNA from wild-type nematodes. The predicted amino acid sequence is exactly as predicted for Y39E4B.9.

bre-4(*ye13*) was placed in *trans* to *unc-35*(*e259*) *dpy-5*(*e61*). 29/32 Dpy nonUncs-segregated *bre-4*(*ye13*) indicated *bre-4* was near and to the right of *unc-35*. Bre nonDpy recombinants from *bre-4*(*ye13*) *dpy-5*(*e61*)/CB4856 heterozygotes placed *bre-4* to the right of the single nucleotide polymorphism Y71G12B.60148. Dpy recombinants from heterozygotes *unc-35*(*e259*) *dpy-5*(*e61*),*bre-4*(*ye43*) (*bre-4* allele in the Hawaiian background) placed *bre-4* to the left of single nucleotide polymorphism Y92H12A.28829. The 150-kilobase region between these boundary single nucleotide polymorphisms contains 20 predicted ORFs. Sequencing of cDNA isolated from the *bre-4* alleles revealed changes in the enzymatically characterized glycosyltransferase Y73E7A.7 (8). cDNA rescue of a *bre-4* mutant confirmed the identity of *bre-4*. A complete cDNA was assembled as for the other *bre* genes (including using spliced leader 1-specific primers). The predicted protein sequence corresponds exactly as predicted for ORF Y73E7A.7.

bre-3(*ye28*) was placed in *trans* to *sma-2*(*e502*) *ced-7*(*n1892*) *unc-69*(*e587*). Of 25 Sma non-Unc recombinants, 12 were *sma-2*(*e502*) *bre-3*(*ye28*) *ced-7*(+), 12 were *sma-2*(*e502*) *bre-3*(+) *ced-7*(*n1892*), and 1 was *sma-2*(*e502*) *bre-3*(+) *ced-7*(+). These data indicated *bre-3* was close and to the left of *ced-7*. We co-injected *bre-3*(*ye28*) animals with cosmids in this region and plasmid pRF4, which encodes the dominant *rol-6* marker, and found that 2/2 stably transmitting lines containing cosmid B0464 were rescued to toxin susceptibility. Subcloning and injection of different regions within B0464 indicated that the rescuing activity (8/8 lines) was conferred by a 4.3-kilobase *Fsp*I-*Nhe*I fragment containing nucleotides 13,437–17,729. This fragment contains a single predicted ORF B0464.4 and an additional ~800 bases at both the 5′ and 3′ ends of the ORF. Sequencing of genomic DNA and cDNA from three *bre-3* mutant alleles indicated these alleles contained alterations in this ORF, confirming the identity of the gene. A complete wild-type cDNA was assembled and sequenced from two overlapping fragments generated 1) at the 5′ end by using random-primed total cDNA, a spliced leader 1 primer, an internal primer, and PCR and 2) at the 3′ end by using a λ phage cDNA library (generously provided by Robert Barstead), a 3′ end-specific vector primer, an internal primer, and PCR. The predicted protein sequence from this assemble cDNA corresponds exactly to that of the predicted B0464.4 ORF. Transmembrane domains in BRE-3 and other related proteins were predicted using TMHMM2.0.

Intestine-specific Expression and Rescue—pBluescript (KS) (Stratagene) was used as the vector backbone for test constructs. The *cpr-1* promoter (9), corresponding to nucleotides 9,541–11,581 of cosmid C52E4, was amplified from *C. elegans* genomic DNA and cloned upstream of *bre* gene coding regions (for *bre-2*, *bre-3*, and *bre-5*, respectively, nucleotides 66,617–69,835 of Y39E4B, 14,285–16,925 of B0464, and 22,694–23,958 of cosmid T12G3; the *bre-4* coding region was amplified from *C. elegans* cDNA). Coding regions were fused in-frame to a double HA tag at their 3′ ends, with a downstream transcription termination region derived from the *bre-5* locus (T12G3 nucleotides 21,578–22,699), except in the *bre-3* construct, which contained the *bre-3* transcription termination region downstream of the HA tag (B0464 nucleotides 13,437–14,287).

Constructs were co-injected with the dominant *rol-6* marker (pRF4) into mutant animals to obtain stably transmitting lines carrying extrachromosomal arrays. Eggs laid by Roller (or N2 control) hermaphrodites were placed on modified ENG agar plates containing 1 mM isopropyl β-D-thiogalactoside and 50 μg/ml carbenicillin and spread with *E. coli* cells expressing Cry5B (5) or with vector-only cells. Eggs were allowed to hatch and grow for 72 h, and resistance was scored based on the ability of larvae to progress to adulthood during the course of the experiment. For each construct tested at least two independent transgenic lines were assayed, and all lines possessed the reported phenotypes with >90% penetrance, with ~20 rollers and 20 non-rollers tested per line.

Fixation and Staining for Immunofluorescence Microscopy—The fixation procedure was a modification of that of Finney and Ruvkun (10). A mixed population of a rescuing line expressing *cpr-1*::*bre-5*::2XHA or cpr-1::*bre-3*::2XHA was suspended in phosphate-buffered saline, 25% methanol, 3% formaldehyde and flash-frozen. After thawing under tap water worms were incubated on ice for 30 min followed by 2 washes in 50 mM sodium borate buffer, 0.01% Triton X-100, pH 9.5 (BO_3T). Worms were then incubated for 20 min in BO_3T plus 10 mM dithiothreitol followed by another wash in BO_3T and a 20-min incubation in BO_3T plus 0.3% hydrogen peroxide. Fixed and permeabilized worms were then washed in phosphate-buffered saline, 0.5% bovine serum albumin, 0.05% Triton X-100, 0.02% sodium azide (antibody buffer). For staining, all steps were carried out in antibody buffer using monoclonal anti-HA antibodies (16B12, Covance), and a final wash included 4′,6-diamidino-2-phenylindole for DNA staining.

Assays for Endocytosis of Rhodamine-labeled Cry5B—Cry5B protoxin was purified as described previously (6) and labeled using a 5-fold molar excess of NHS-rhodamine. Trypsinization was carried out at a 1:200 (trypsin:toxin) mass ratio for 2 h at 25 °C. Trypsinized toxin was still toxic but 5–10-fold less based on growth assays. We used Edman degradation sequencing and determined that trypsin removes the first 170 amino acids, which includes predicted α helix 3 but not helix 4. Consistent with our results studies with insecticidal Cry toxins show that α helices 4 and 5 are critical for toxin function and that helices 1–3 may contribute some activity (11, 12). Based on the apparent molecular weight of trypsinized toxin, trypsin is predicted to remove most of the carboxyl-terminal protoxin domain. Animals (L4 stage) were incubated in M9 medium containing 50 μg/ml protoxin or 120 μg/ml trypsinized toxin for 3 h and washed 3 times in M9 before imaging. For wash-out experiments, *bre-4*(*ye13*) and *bre-5*(*ye17*) animals fed rhodamine toxin for a few hours in wells were pipetted onto standard NG plates and examined every few minutes on an Olympus green fluorescent protein dissecting scope fitted with a rhodamine filter.

Single Mutant Versus Double Mutant Resistance Assays—All double mutants were constructed and confirmed using a similar strategy. For example, the *bre-4*(*ye13*),*bre-5*(*ye17*) double mutant was created by mating *bre-4*(*ye13*) males into the strain *dpy-5*(*e61*),*bre-5*(*ye17*) (*bre-4* and *dpy-5* are both on chromosome I). NonDpy F1 progeny were cloned, and in the F2 nonDpy Cry5B-resistant animals were picked that segregated F3 Dpy (*bre-4*(*y13*),*dpy-5*(*e61*),*bre-5*(*ye17*)). NonDpy F3 were cloned out, and those that did not segregate Dpy were selected as *bre-4*(*ye13*),*bre-5*(*ye17*) putative doubles. The presence of both mutations was verified by sequencing DNA.

To measure resistance of single and double mutants we used the

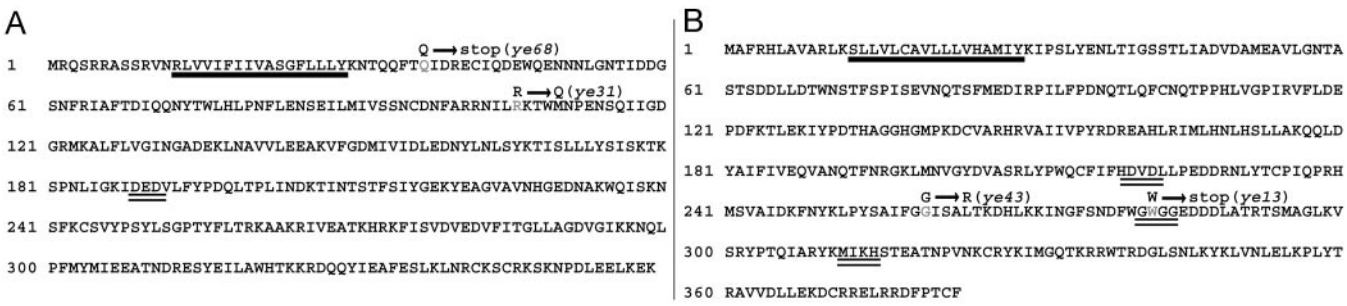

FIG. 1. ***bre-2* and *bre-4* genes encode a putative and a known glycosyltransferase, respectively.** The protein sequences shown are assembled by translating *bre-2* and *bre-4* cDNAs. BRE-2 (*A*) and BRE-4 (*B*) sequences are shown. *Solid underlines* denote the predicted membrane-spanning regions. *Double underlines* indicate conserved motifs known to play key catalytic roles in related glycosyltransferases. Mutations observed in various alleles are also shown.

Cry5B-related crystal protein Cry14A (5, 6). Cry14A is required for these studies since *bre-2*, *-3*, *-4*, and *-5* are very highly resistant to Cry5B (*e.g.* they are not killed by even 1 mg/ml Cry5B) but are only moderately resistant to Cry14A. Cry14A was produced in *E. coli* and incorporated into the assay diet as described (6). Assays were carried out in 24-well plates, with each well containing 400 μl of S medium with 30 μg/ml each of tetracycline and chloramphenicol and ~60 newly hatched L1 animals, with a final bacterial density of A_{600} = 0.25. After a 60-h exposure at 20 °C, animals were mounted and imaged, and area measurements were made outlining the nematodes by hand and processing with NIH Image.

bre-5 Rescue of the Drosophila brainiac Mutant—The *bre-5* cDNA was cloned into pUAST (13) and used to transform *yellow*, *white* embryos using standard techniques. Two independent lines containing the transgene on chromosome II were used for *brn* complementation assays, which were carried out as described in Müller *et al.* (14). Briefly, male UAS-*bre-5* transformants were crossed to females of the genotype *forked*, $brn^{1.6P6}$/FM6,w^1, actin-, heat shock- or *armadillo*-GAL4,w^+. Rescue was scored based on the ability of *forked* male progeny to hatch. The GAL4 drivers used for this experiment carry the Bloomington stock numbers 1560 (*armadillo*-GAL4), 2077 (heat-shock GAL4), and 4414 (actin-GAL4). These lines were generously provided by Markus Noll and Erich Frei (Institute of Molecular Biology, University of Zürich). Flies carrying UAS-*brainiac* and UAS-β3GnT-IV (a human *N*-acetylglucosaminyltransferase) were used as positive and negative controls, respectively.

bre-5-dependent N-Acetylglucosaminyltransferase Assays—N2 and *bre-5*(*ye17*) mixed larval-staged animals were reared on ENG medium (standard NG with an additional 2.5 g of peptone and 1 g of yeast extract per liter), washed, flash-frozen in 50 mM Tris, pH 7.5, as a 0.2-ml pellet, and stored at −70 °C. Samples were sonicated on ice in 0.05% Triton-X-100, 50 mM Tris and centrifuged at 15,000 × *g* for 10 min to obtain a clarified supernatant. The assay mixture (25 μl) contained 50 mM cacodylate, pH 6.5, 0.1% taurodeoxycholate, 10 mM $MnCl_2$, 1 μCi of UDP-[6-^{3}H]GlcNAc (39.7 Ci/mmol, PerkinElmer Life Sciences),10 mM Manβ1,4Glc-pNP (custom synthesis of Toronto Research, ON, Canada), and 8 μg of extracted protein (N2 or *bre-5*) as the enzyme source. After incubation at 25 °C for 7 h, the reaction products were diluted with 0.5 ml of 0.5 M NaCl and applied to a Sep-Pak C18 cartridge (100 mg; Waters). After washing the cartridge with 25 ml of water, the products were eluted with 50% methanol, dried, and counted by liquid scintillation.

RESULTS

To better understand how resistance to pore-forming Cry toxins develops, we used single-nucleotide polymorphisms to map the *bre-2* and *bre-4* genes to small molecular intervals of ≤20 genes, each of which contained a glycosyltransferase gene. Given that the previously identified *bre-5* was shown to encode a glycosyltransferase gene (6), we sequenced cDNAs isolated from *bre-2* and *bre-4* mutant alleles and found that multiple alleles of each were associated with point mutations in the glycosyltransferase genes Y39E4B.9 and Y73E7A.7, respectively (Fig. 1, *A* and *B*). At least one allele for each gene (*bre-2*(*ye68*) and *bre-4*(*ye13*)) is predicted to eliminate protein function by truncation of the putative glycosyltransferase catalytic domain, indicating that Cry5B resistance is the null phenotype for each. That RNA-mediated interference (RNAi) of *bre-2* recapitulates Cry5B resistance supports this (data not shown). Rescue data (see below) confirm that these glycosyltransferase genes encode *bre-2* and *bre-4*. Cloning and characterization of *bre-3*, which encodes an unusual glycosyltransferase of striking resemblance to a known glycosyltransferase in *Drosophila*, is dealt with below.

bre-2 encodes a putative family 31 β1,3-glycosyltransferase with a single amino-terminal membrane-spanning domain and an appropriately positioned D*X*D motif that is conserved in this enzyme family and is thought to be critical for catalysis (Fig. 1*A*). In general, BRE-2 shares ~25% amino acid identity in its putative enzymatic domain with other β1,3-glycosyltransferases such as with BRE-5 (a β1,3-GlcNAc transferase; see Griffitts *et al.* (6) and below) and insect and mammalian β1,3-Gal and GlcNAc-transferases. *bre-2* is also a member of a more closely related and expanded subfamily of *C. elegans* β1,3-glycosyltransferase genes that includes T09F5.1, F14B6.6, T09E11.10, E03H4.11, C54C8.3, F14B6.4, C47F8.3, C47F8.5, C47F8.6, and T15D6.5 and to which it shares ~40% identity in its glycosyltransferase domain. To our knowledge, *bre-2* is the first one of these family members to be characterized genetically and to be demonstrated to have a loss of function phenotype.

bre-4 is predicted to encode a member of the family 7 glycosyltransferases, of which β1,4-galactosyltransferases are the best characterized (15). BRE-4 is predicted to contain a single membrane-spanning region near its amino terminus as well as conserved residues involved in catalysis (Fig. 1*B*). BRE-4 has been previously characterized biochemically as a UDP-GalNAc:GlcNAc β1,4-*N*-acetylgalactosaminyltransferase (Ceβ4 GalNAcT) (8). BRE-4/Ceβ4GalNAcT specifically synthesizes the GalNAcβ1,4GlcNAc sequence commonly found on glycoproteins and glycolipids. No mutational analysis of *bre-4* has been previously reported. There are two other related glycosyltransferase genes in the *C. elegans* genome. Blast searches using both BRE-4 full-length protein and its enzymatic domain indicate a best hit and likely homologue in *Drosophila melanogaster* (predicted gene CG8536) and *Anopheles gambiae* (ORF from nucleotides 565443 to 566142, chromosome II). The predicted catalytic domains of these proteins, respectively, share 45 and 51% identity to the BRE-4 catalytic domain.

Expression of bre-5, bre-2, and bre-4 in the Intestine Is Required for Cry Protein Intoxication—Based on the fact that Cry toxins appear to specifically attack the gut cells of insects and nematodes, we predicted that *bre-2* and *bre-4* glycosyltransferase genes should be required in intestinal cells for Cry protein intoxication. To demonstrate this intestinal requirement while simultaneously confirming the correct identities of *bre-2* and *bre-4*, we placed wild-type *bre-2* genomic coding re-

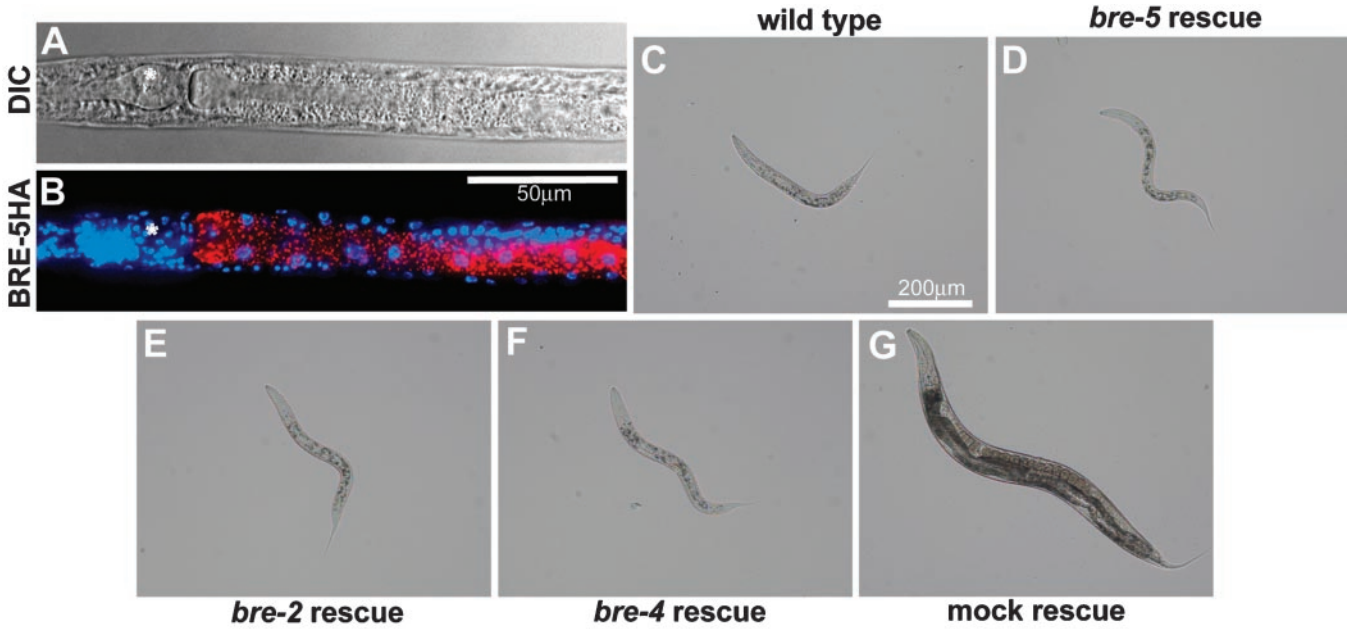

FIG. 2. **Intestine-specific expression of *bre-5*, *bre-2*, and *bre-4* are sufficient for Cry5B intoxication and confirm cloning of *bre-2* and *bre-4*.** *A* and *B*, differential interference contrast (*DIC*) and fluorescence microscopy verifies intestine-specific expression of the *cpr-1* promoter driving BRE-5::HA (BRE-5::HA is in *red*, and 4′,6-diamidino-2-phenylindole is in *blue*). *Asterisks* denote pharyngeal terminal bulb. *C–G*, images of animals after 72 h of growth on Cry5B, all at the same magnification and oriented with heads pointing *upper left*. *C–F* are toxin-sensitive; *G* is toxin-resistant. *C*, N2 on Cry5B. *D*, *bre-5*(*ye17*) transformed with *cpr-1*::*bre-5* on Cry5B. *E*, *bre-2*(*ye31*) transformed with *cpr-1*::*bre-2* on Cry5B. *F*, *bre-4 (ye13)* transformed with *cpr-1*::*bre-4* on Cry5B. *G*, *bre-5*(*ye17*) transformed with *cpr-1*::*bre-2* on Cry5B; *bre-2* fails to rescue the *bre-5* mutant.

gion and *bre-4* cDNA downstream of the intestinal-specific *C. elegans* promoter from the *cpr-1* gene (9). As a control using the previously cloned *bre-5* gene we also made a *cpr-1*::*bre-5* construct in which a 2× HA tag was fused in-frame at the carboxyl terminus of *bre-5*. Intestine-specific expression of the *cpr-1* promoter was confirmed by HA antibody staining of a stably transmitting line transformed with the *cpr-1*::*bre-5*:HA construct (Fig. 2, *A* and *B*). This construct was fully capable of restoring Cry5B susceptibility to a *bre-5*(*ye17*) mutant (Fig. 2, *C* and *D*), demonstrating that intestinal expression of wild-type *bre-5* in an animal otherwise lacking *bre-5* is sufficient for Cry protein intoxication and providing a complementary result to our mosaic experiments (6).

We similarly found that intestinal-specific expression of wild-type *bre-2* in a *bre-2*(*ye31*) mutant background and intestinal-specific expression of wild-type *bre-4* in a *bre-4*(*ye13*) mutant background fully restored Cry toxin susceptibility to each (Fig. 2, *E* and *F*). Thus, like *bre-5*, both glycosyltransferase genes act in the intestine in their mediation of Cry protein intoxication. The fact that the *bre-2* gene complements a *bre-2* mutant and that the *bre-4* gene complements a *bre-4* mutant also confirms the identities of these genes. To demonstrate specificity of rescue, we transformed a *bre-5*(*ye17*) mutant with the wild-type *cpr-1*::*bre-2* construct that fully rescues a *bre-2* mutant. No rescue of the *bre-5* mutant by the *bre-2* gene was seen (Fig. 2*G*) despite the fact that both are members of the extended β1,3-glycosyltransferase family.

Toxin Uptake in C. elegans Is Associated with Active Toxin and Requires bre-2, bre-4, and bre-3—We previously demonstrated that *bre-5* is required for Cry5B toxin to interact with intestinal cells since *bre-5*(*ye17*) mutant animals fail to endocytose Cry5B protoxin labeled with rhodamine (6). However, because Cry proteins are produced as large (protoxin) precursors that are proteolytically activated, it was not demonstrated that toxin molecule taken up was the active toxin fragment or that the toxin had to be functional to be taken up. To extend these results, we fed *C. elegans* rhodamine-labeled Cry5B that was full-length (protoxin), proteolytically activated with trypsin (Fig. 3*A*), or boiled to inactive it. As reported (6), full-length protoxin is readily taken up by intestinal cells into autofluorescent gut granules, which are reported to be the secondary lysosome (Fig. 3, *B* and *C*). We found that trypsin-activated toxin is also endocytosed by intestinal cells (Fig. 3*D*). In contrast, inactivated toxin is not taken up (Fig. 3*E*). Thus, toxin uptake into intestinal cells is associated with activated, functional toxin.

In contrast to wild-type animals, *bre-2*(*ye31*) *bre-4*(*ye13*), and *bre-3*(*ye28*) (*bre-3* is discussed further below), mutant animals do not endocytose either functional protoxin or trypsin-activated toxin into their intestinal cells (Fig. 3, *F–K*). Rhodamine toxin seen in the intestinal lumen is not stably associated with the intestinal membrane since *bre-4*(*ye13*) and *bre-5*(*ye15*) mutant animals fed a pulse of rhodamine toxin flush the labeled toxin from their lumen within 5–10 min; no residual binding was detected (Fig. 3*L*, shown for *bre-4*(*ye13*)). These results suggest that the glycosyltransferase genes make a component required for toxin to interact with the plasma membrane of the intestine. An alternative explanation, that these mutants are all defective in general endocytosis by the intestine is unlikely since all the mutants grow at relatively normal rates and have good brood production and because we have already demonstrated that *bre-5*(*ye17*) mutants are not obviously defective in endocytosis (6).

bre-3 Encodes the C. elegans Homologue of Drosophila Egghead (egh)—Although the identities of *bre-2* and *bre-4* were not obviously linked to each other or *bre-5* apart from the fact all are glycosyltransferase genes, the identification of *bre-3* is striking given the identity of *bre-5*. *bre-3* was cloned by genetic mapping followed by rescue with cosmid B0464 and then rescue with the single ORF B0464.4. *bre-3* encodes a putative glycosyltransferase belonging to family 2 β transferases (Fig. 4*A*) and includes aspartate residues conserved in this family. Many enzymes in this family, including BRE-3, also contain a characteristic QRXRW motif that is important for transferase function (16). Based on sequencing mutant alleles, Cry5B resistance is the null phenotype for *bre-3* since several known alleles

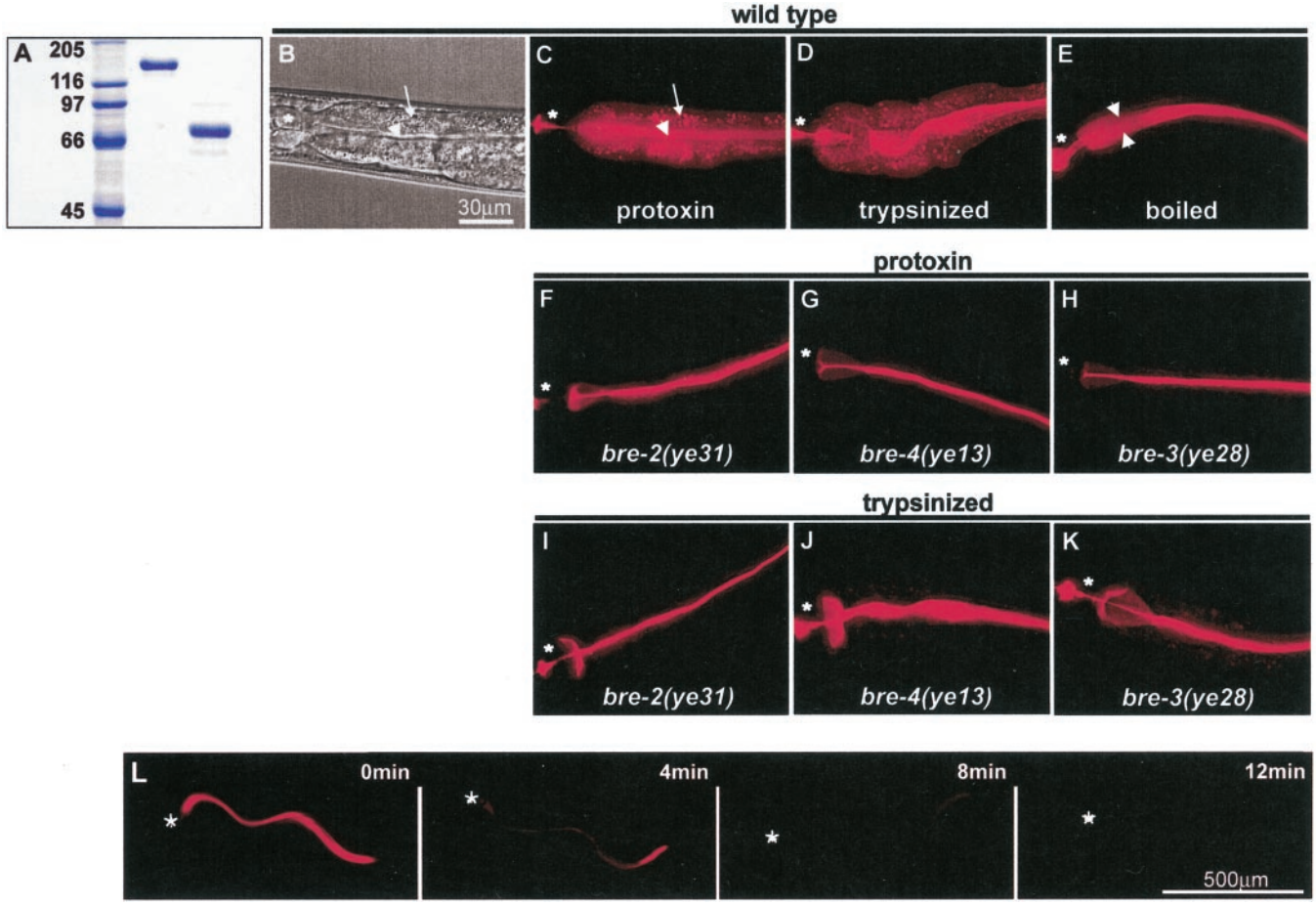

FIG. 3. **Uptake of Cry5B toxin into intestinal cells is associated with active toxin and is defective in *bre-2*, *bre-4*, and *bre-3* mutants.** *A*, protoxin and trypsinized toxin visualized by SDS-PAGE. *Lane 1*, molecular weight markers; *lane 2*, purified Cry5B protoxin; *lane 3*, purified Cry5B toxin processed with trypsin. *C–L*, fluorescent images of animals fed rhodamine-labeled toxins. *B* and *C*, matching differential interference contrast and fluorescent image of wild-type animal fed labeled protoxin. *C*, wild-type intestines take up labeled protoxin. *D*, wild-type intestines take up trypsinized toxin. *E*, wild-type intestines do not take up inactivated protoxin. *F–H*, *bre-2*(*ye31*), *bre-4*(*ye13*), and *bre-3*(*ye28*) mutant animals fail to take up active, labeled protoxin. *I–K*, *bre-2*(*ye31*), *bre-4*(*ye13*), and *bre-3*(*ye28*) mutant animals fail to take up active, labeled trypsinized toxin. *L*, progressive and rapid loss of rhodamine signal in a *bre-4*(*ye13*) animal fed labeled toxin in a well and then transferred to an *E. coli* lawn (wash-out experiment). *Asterisks* denote pharyngeal terminal bulb; *arrows* denote intracellular intestinal granules; *arrowheads* indicate intestinal lumen. *B-K* are at the same magnification.

(*ye9* premature stop; *ye28* internal deletion) are predicted to eliminate protein function. RNA-mediated interference of *bre-3* recreated Cry5B resistance, confirming this conclusion. *bre-3* mutant animals appear otherwise healthy.

BRE-3 has a very high level of sequence identity to *Drosophila* EGGHEAD (EGH) (60% amino acid identity over their entire lengths, Fig. 4*A*). BRE-3 also is highly similar to a predicted EGH-like enzyme in *A. gambiae* (59% identity; Fig. 4*A*). BRE-3 and EGH differ in length by only two amino acids at the very carboxyl terminus, and both are unique in their respective genomes.

EGH is known to specifically catalyze the transfer of mannose to glucose with a β1,4-glycosidic linkage (17). BRE-3 and EGH also display low, but detectable, levels of sequence identity to processive glycosyltransferases from bacteria (*e.g.* NodC (Fig. 4*A*) and cellulose synthase, not shown), suggesting that perhaps the invertebrate family evolved from the bacterial family. Consistent with this hypothesis, analysis of potential transmembrane regions for BRE-3, EGH, NodC, and other proteins of this class indicate striking conservation of structure not evident in sequence alignments. These enzymes are predicted to be anchored in the membrane at both the amino and carboxyl termini, with the enzymatic domain (amino acids 80–330 for BRE-3) all predicted to sit in the cytosol (Fig. 4, *B–D*). BRE-3/EGH has no detectable amino acid sequence similarity to proteins from vertebrates.

We expressed 2× HA-tagged BRE-3 under the control of the *cpr-1* intestinal promoter and analyzed its localization in intestinal cells (Fig. 5*A*). BRE-3::HA protein localizes to a punctate intracellular compartment. No detectable protein was seen at the plasma membrane. These results are consistent with BRE-3 acting as a glycosyltransferase that intracellularly modifies a Cry toxin receptor and not, despite the presence of multiple putative TM domains, as a direct receptor for Cry toxin at the plasma membrane. As with *bre-2*, *bre-4*, and *bre-5*, intestinal-specific expression of *bre-3* was sufficient to restore susceptibility to toxin in a *bre-3*(*ye28*) mutant background (Fig. 5, *B* and *C*).

The Four bre Glycosyltransferase Genes Function in a Single Pathway—The identity of the *bre-3* gene as the *C. elegans* putative *egh* homologue is striking given the previous identification of *bre-5* as having some sequence identity to *Drosophila brainiac* (*brn*). *brn* and *egh* display similar neurogenic phenotypes in *Drosophila*, act in the same genetic pathway in epithelial morphogenesis of *Drosophila* follicle cells, and have been shown to catalyze consecutive glycosylation reactions (14, 17–20). Based on enzymology, it is predicted that *egh* and *brn* function to synthesize *Drosophila* glycosphingolipids, although

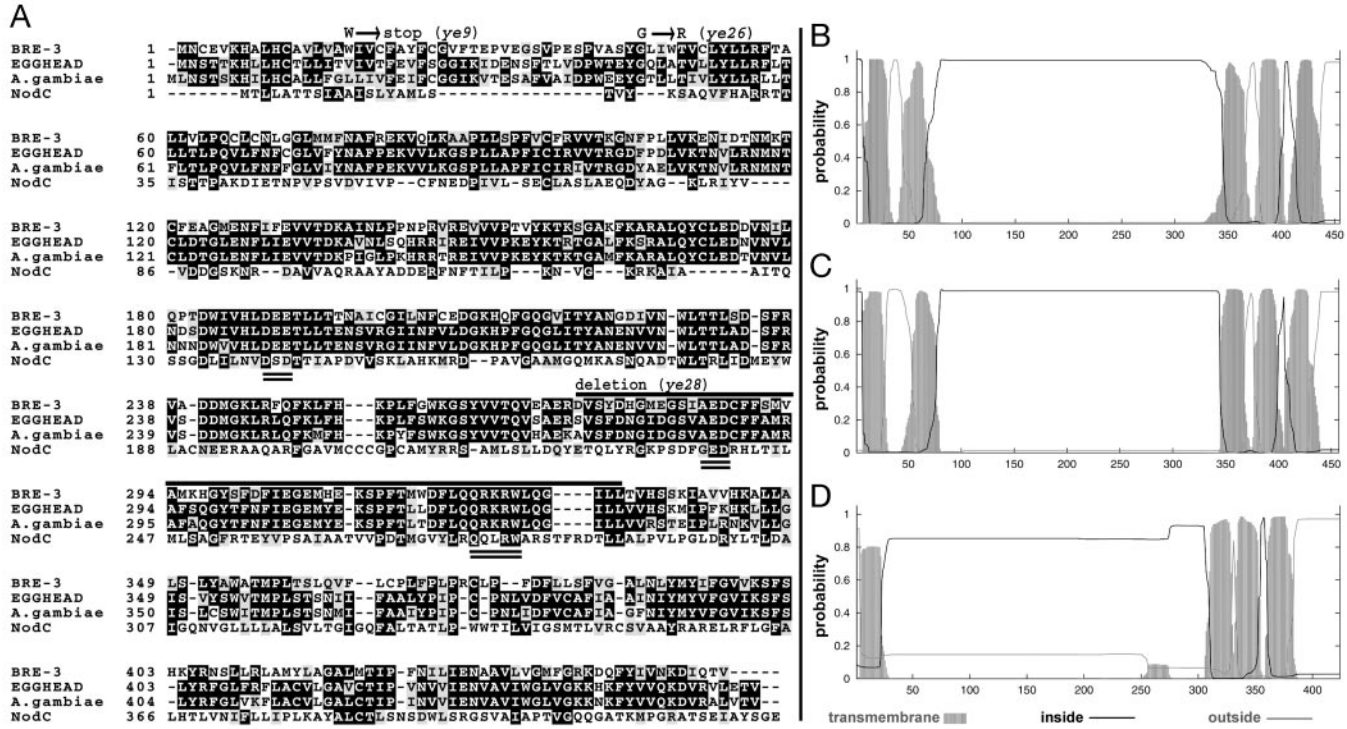

FIG. 4. ***bre-3* encodes the *C. elegans* EGGHEAD homologue.** *A*, the protein sequence shown is assembled by translating *bre-3* cDNA. Alignments to EGGHEAD from *D. melanogaster*, the EGGHEAD homologue in *A. gambiae*, and the NodC protein from *Rhizobium leguminosarum* are shown. Identical residues between BRE-3, EGH, and *A. gambiae* EGH are *highlighted* in *black boxes*. *Double underlines* indicate conserved motifs known to play key catalytic roles in related glycosyltransferases. Mutations observed in various *bre-3* alleles are also shown. The *bre-3*(*ye28*) allele is due to a DNA deletion corresponding to residues beneath the *solid line* and leads to a frameshift over the remainder of the coding region. *B–D*, predicted transmembrane regions and topology of BRE-3, EGH, and NodC, respectively.

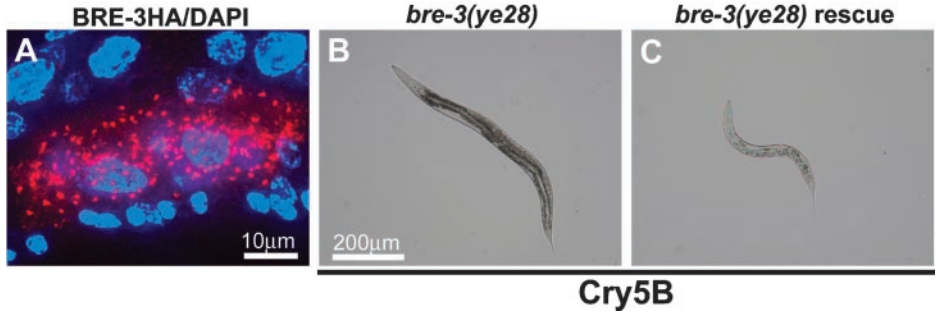

FIG. 5. **BRE-3 is in cytoplasmic puncta and is required in the intestine for intoxication by Cry5B.** *A*, double staining (HA is in *red*; 4′,6-diamidino-2-phenylindole (*DAPI*) is in *blue*) of HA-tagged BRE-3 driven by the *cpr-1* promoter. BRE-3 localizes to cytoplasmic puncta and is not seen at the plasma membrane. *B*, a resistant *bre-3*(*ye28*) animal on Cry5B toxin. *C*, *bre-3*(*ye28*) transformed with *cpr-1*::*bre-3*::HA construct demonstrating intestinal-specific rescue.

glycosphingolipid defects in the *Drosophila* mutants have yet to be shown. The similarities of *bre-3* to *egh* and *bre-5* to *brn* suggest that Cry protein intoxication in *C. elegans* uses the same pathway as that required for epithelial morphogenesis in *Drosophila*. To test whether *bre-3* and *bre-5* act in a common pathway for Cry5B intoxication, we examined the dose response of animals homozygous for the molecular null alleles *bre-3*(*ye28*), *bre-5*(*ye17*), and the double mutant *bre-3*(*ye28*),*bre-5*(*ye17*) to the Cry5B-related crystal toxin Cry14A (see "Experimental Procedures"). Each single and the double mutant shows a similar level of Cry toxin resistance (Fig. 6). These data indicate that, like *egh* and *brn*, *bre-3* and *bre-5* function in a single, common genetic pathway.

Given their similar phenotypes and identification as glycosyltransferase genes, we hypothesized that *bre-2* and *bre-4* also function in the same pathway as *bre-3*-*bre-5*. To test this hypothesis double mutants were made between *bre-4*(*ye13*) and each of the other mutants. The lesions in the alleles used are predicted to result in deletions of required catalytic residues, except *bre-2*(*ye31*), which results in a single amino acid change at a highly conserved residue. Resistance to Cry14A toxin was extensively quantitated in all single and double mutants at three doses (Fig. 7). Resistance levels of each of the four single mutants are similar, and at any given dose, resistance levels of all double mutant strains were not statistically different ($p \geq 0.2$) in pairwise comparisons with the reference *bre-4*(*ye13*) single mutant strain. Thus *bre-2*, *bre-3*, *bre-4*, and *bre-5* form an extended single genetic pathway for Cry toxin resistance.

bre-5 Is the Functional Homologue of Drosophila brn and Provides the Major UDP-GlcNAc:Man N-Acetylglucosaminyltransferase Activity in C. elegans—Although BRE-3 and EGH share high levels of amino acid identity and are unique in their respective genomes, BRE-5 and BRN are ~29% identical (37% in their predicted catalytic domains) and overall belong to the larger family of β1,3 glycosyltransferases, of which there are many predicted in both the *C. elegans* and *Drosophila* genomes

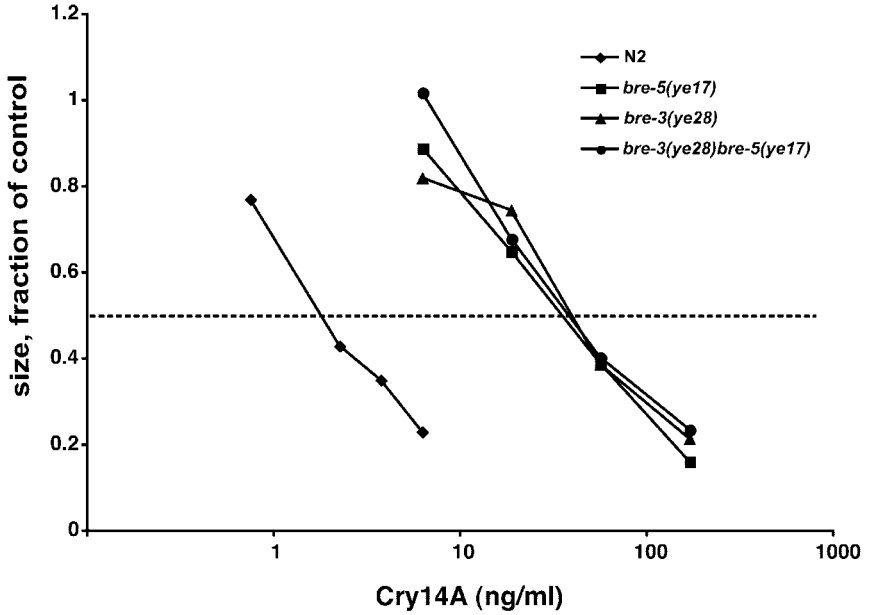

FIG. 6. ***bre-3* and *bre-5* function in a single genetic pathway.** Cry14A dose-response curves for N2, *bre-3*(*ye28*), *bre-5*(*ye17*), and *bre-3*(*ye28*),*bre-5*(*ye17*) animals. The strains (as shown) were exposed to four doses of Cry14A. Data shown represent the average sizes attained over the experiment as a fraction of controls fed non-toxic *E. coli*. Each data point represents the average size of 8–20 animals (on average 12).

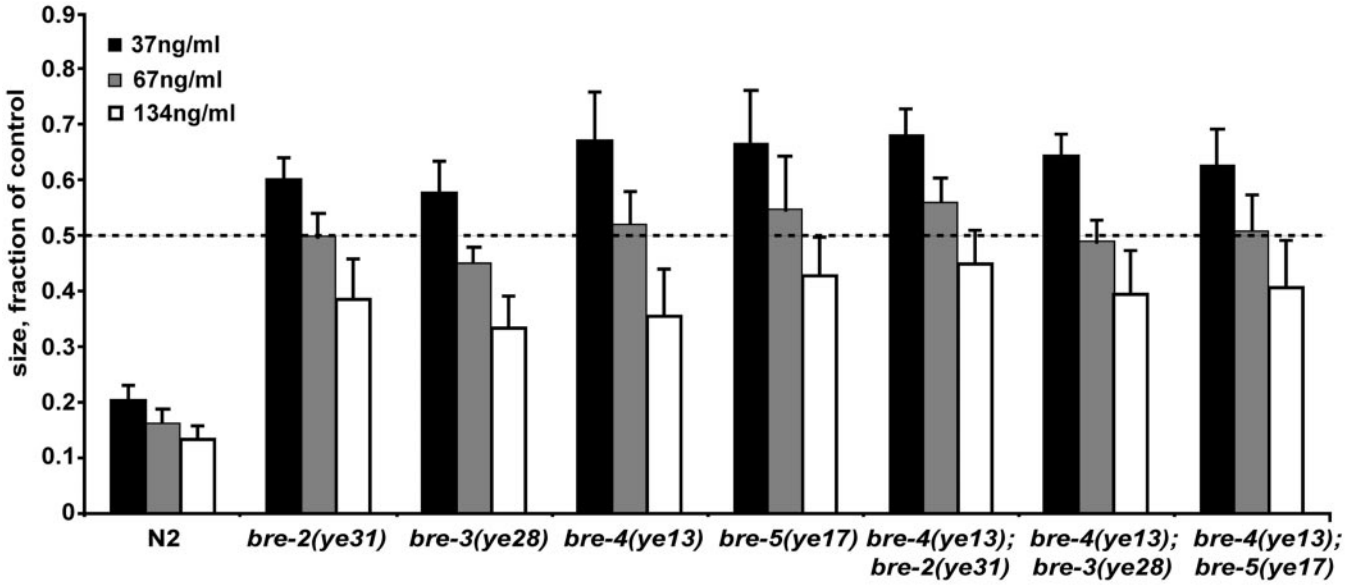

FIG. 7. **All four *bre* genes function in a common genetic pathway.** The strains (as shown) were exposed to 37 ng/ml (*black bars*), 67 ng/ml (*gray bars*), and 134 ng/ml (*white bars*) of Cry14A. Data shown represent the average sizes attained over the experiment as a fraction of controls fed non-toxic *E. coli*. *Error bars* represent the S.D. from the mean of averages over three independent experiments, each employing 20–30 animals per condition.

(including *bre-2*; see above). To therefore test whether BRE-5 and BRN are true homologues, we drove expression of a wild-type copy of *C. elegans bre-5* cDNA in *Drosophila* mutant for *brn*. Using three different promoters, we found that both *brn* and *bre-5* could complement the lethality associated with *Drosophila* homozygous for the $brn^{1.6P6}$ (21) mutation (Table I). In contrast, a human β1,3-*N*-acetylglucosaminyltransferase was not able to rescue *brn* lethality. Thus, BRE-5 is the *C. elegans* homologue of *Drosophila* BRN.

The functional homology of *bre-5* and *brn* predicts that *bre-5* should encode *N*-acetylglucosaminyltransferase (GlcNAc-transferase) activity. Enzymatic experiments performed with extracts from wild-type animals demonstrate that *C. elegans* has GlcNAc-transferase activity when Manβ1,4GlcpNP is used as the acceptor (the preferred acceptor for BRN) (Fig. 8). This activity returns almost to background levels when extracts are prepared from *bre-5*(*ye17*) homozygous animals instead of wild type (Fig. 8). These data indicate that BRE-5 is a UDP-GlcNAc:Man *N*-acetylglucosaminyltransferase and that BRE-5 provides the majority of such activity in *C. elegans*.

TABLE I
Complementation of D. melanogaster $brn^{1.6P6}$ by bre-5

UAS transgenic Line	GAL4 transgenic lines		
	Actin GAL4	Armadillo GAL4	Heat shock GAL4
	% rescue[a]		
brn	87 (47)	97 (403)	85 (124)
bre-5 (line 1)	25 (231)	20 (289)	25 (146)
bre-5 (line 2)	78 (117)	57 (266)	35 (166)
Human β3GnT-IV	0 (90)	0 (116)	0 (130)

[a] Rescue is given as percentage of hatched *forked, $brn^{1.6P6}$* males relative to the number of hatched males carrying the balancer chromosome FM6,w^1. The number of males used for the calculation is indicated in parentheses.

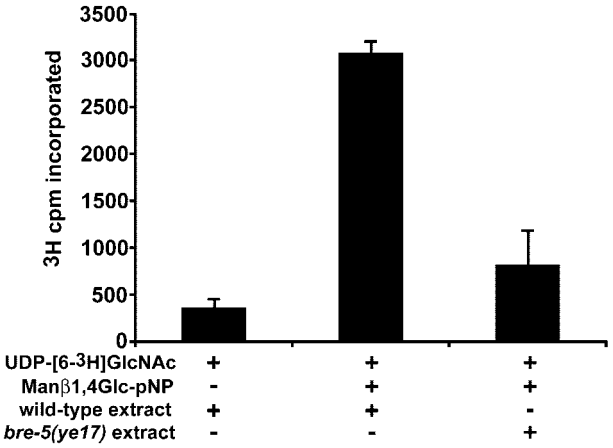

FIG. 8. ***bre-5* encodes for the majority of UDP-GlcNAc:Man *N*-acetylglucosaminyltransferase activity in *C. elegans*.** Crude extracts were tested for the ability to transfer radiolabeled GlcNAc from UDP-[6-^{3}H]GlcNAc to the acceptor sugar Manβ1,4Glc-pNP.

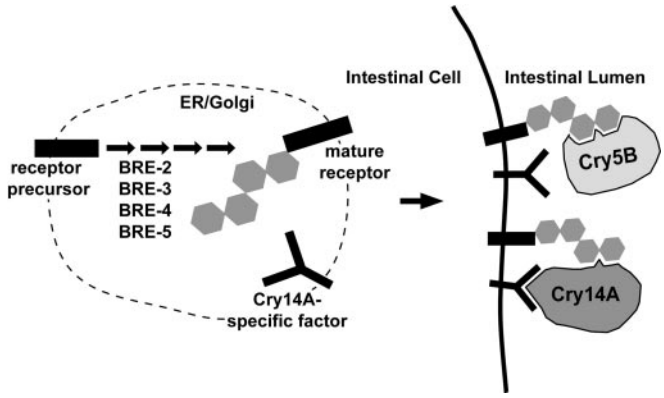

FIG. 9. **An oligosaccharide receptor-based model for the *bre* genes in intestinal cells that includes the overlapping, but non-identical requirements of the *bre* genes for Cry5B and Cry14A intoxication.**

DISCUSSION

Characterization of the four major genes that mutate to confer resistance to pore-forming crystal toxin in *C. elegans*, *bre-2*, *bre-3*, *bre-4*, and *bre-5*, has led to the identification of a single glycosylation pathway. Components of this same pathway are active in insects. The *egh* and *brn* glycosyltransferase genes required for epithelial development in *Drosophila* correspond to the *bre-3* and *bre-5* Cry5B resistance genes in *C. elegans*. We demonstrate that *bre-5* and *brn* are functional homologues. It is likely that *bre-3/egh* are functional homologues as well since *bre-3* and *egh* are unique in their respective genomes, share a high level of amino acid identity, and function in a common pathway with *bre-5/brn*.

Given the lack of a vertebrate *bre-3/egh* homologue, the functional conservation of *bre-5* and *brn*, and the fact that *bre-5* and *brn* are more similar to each other than they are to mammalian β1,3-glycosyltransferases, it appears that the *bre-3/egh-bre-5/brn* pathway is an invertebrate-specific glycosylation pathway. Enzymatically, *egh* and *brn* catalyze consecutive steps that are consistent with the production of glycosphingolipids, although this has yet to be demonstrated *in vivo*. Our identification of two additional genes, *bre-2* and *bre-4*, extends this pathway. The *bre-4* gene product has been previously demonstrated to have glycosyltransferase activity (8); the *bre-2* gene product has not, and thus, BRE-2 should be regarded as a putative glycosyltransferase. A likely *bre-4* homologue in *Drosophila* is noted above.

All four *bre* mutants are defective in the uptake of toxin into intestinal cells. Furthermore, chase experiments in live animals indicate that toxin is not significantly associated with intestinal membranes in *bre-4* or *bre-5* mutant animals. These data suggest that the *bre* genes are involved in promoting the interaction of Cry5B and Cry14A crystal toxins with the intestinal plasma membrane. The simplest interpretation of these data is that the *bre* genes synthesize an oligosaccharide that is transported to the apical membrane of the intestine that serves as a receptor for crystal toxins (Fig. 9). Ample precedent exists for bacterial toxins binding to extracellular carbohydrates including aerolysin, cholera toxin, and Shiga toxin. Because loss of any of the *bre* enzymes leads to high levels of resistance to Cry5B, we hypothesize that this receptor is the major receptor for Cry5B. In the case of Cry14A, the reduced level of resistance conferred by the mutants suggests that other receptors can partly compensate in the absence of the *bre* oligosaccharide (Fig. 9). An alternative interpretation of our data is that the *bre* genes do not make a direct receptor for toxin but, rather, influence the membrane or membrane proteins in a way that promotes toxin binding and/or uptake.

This work reiterates and provides striking in *vivo* genetic evidence for the importance of carbohydrates in the interaction of bacterial toxins with their host. Our screen for healthy, robust resistant mutants did not uncover any mechanisms of resistance other than the *bre* glycosylation pathway. Given the number of alleles we isolated for each of these genes, it is likely that the list of strong, healthy Cry5B resistance genes is complete or nearly complete. Our data also suggest that receptor binding of a bacterial pore-forming toxin is a central target for potential therapeutics. For example, the modulation of carbohydrates (*e.g.* using synthetic carbohydrates as competitive inhibitors) could have therapeutic benefits against bacterial toxins that require carbohydrates for interacting with the host. Such a strategy is being explored with Shiga toxin in mammals (22).

Our results have important implications for the biology of Cry toxins and Cry toxin resistance as well. The identification of *bre-2*, *bre-3*, and *bre-4* doubles the number of Cry toxin resistance genes known. That the *bre* genes mutate to resistance for two divergent crystal proteins, Cry5B and Cry14A (~34% identical at the amino acid level) suggests that they will be relevant to the biology of other Cry toxins as well. Significantly, Cry14A has been reported to be toxic to coleopteran insects (23), suggesting that insects could use the same receptor. In addition, some insect resistance to Cry toxin has been has been correlated with changes in complex carbohydrate structures (24, 25). Along with the striking conservation of the *bre-3-bre-5* pathway with the *egh-brn* pathway in insects, these observations suggest our results may be relevant to Cry toxin resistance in insects as well. Unlike the *C. elegans bre* mutants, *Drosophila* mutants in this pathway are often lethal. Thus insects may be less able to tolerate loss of function of this pathway. Perhaps this is part of the reason for the high efficacy of insecticidal Cry toxins even over long periods of selection (26). However, viable/sterile mutant alleles of *egh* and *brn* exist. It is, therefore, possible that mutations in these could contribute to the development of resistance in insects. This contribution could be investigated by analysis of homologous sequences in resistant insects and could provide vital information about how invertebrate pests and disease vectors evolve resistance to Cry toxins and how the emergence of resistance can be minimized.

Acknowledgments—We thank Karen Chien for help with bioassays, Dr. James Posakony and members of the Aroian Laboratory for discussions, and Dr. Robert Barstead for his phage library. Some strains were

provided by the *C. elegans* Genetics Center (funded by the National Institutes of Health National Center for Research Resources).

REFERENCES

1. Gilbert, R. J. (2002) *Cell. Mol. Life Sci.* **59,** 832–844
2. Schmitt, C. K., Meysick, K. C., and O'Brien, A. D. (1999) *Emerg. Infect. Dis.* **5,** 224–234
3. Betz, F. S., Hammond, B. G., and Fuchs, R. L. (2000) *Regul. Toxicol. Pharmacol.* **32,** 156–173
4. Marroquin, L. D., Elyassnia, D., Griffitts, J. S., Feitelson, J. S., and Aroian, R. V. (2000) *Genetics* **155,** 1693–1699
5. Wei, J. Z., Hale, K., Carta, L., Platzer, E., Wong, C., Fang, S. C., and Aroian, R. V. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 2760–2765
6. Griffitts, J. S., Whitacre, J. L., Stevens, D. E., and Aroian, R. V. (2001) *Science* **293,** 860–864
7. Brenner, S. (1974) *Genetics* **77,** 71–94
8. Kawar, Z. S., Van Die, I., and Cummings, R. D. (2002) *J. Biol. Chem.* **277,** 34924–34932
9. Britton, C., McKerrow, J. H., and Johnstone, I. L. (1998) *J. Mol. Biol.* **283,** 15–27
10. Finney, M., and Ruvkun, G. (1990) *Cell* **63,** 895–905
11. Uawithya, P., Tuntitippawan, T., Katzenmeier, G., Panyim, S., and Angsuthanasombat, C. (1998) *Biochem. Mol. Biol. Int.* **44,** 825–832
12. Vachon, V., Prefontaine, G., Coux, F., Rang, C., Marceau, L., Masson, L., Brousseau, R., Frutos, R., Schwartz, J. L., and Laprade, R. (2002) *Biochemistry* **41,** 6178–6184
13. Brand, A. H., and Perrimon, N. (1993) *Development* **118,** 401–415
14. Müller, R., Altmann, F., Zhou, D., and Hennet, T. (2002) *J. Biol. Chem.* **277,** 32417–32420
15. Ramakrishnan, B., Boeggeman, E., and Qasba, P. K. (2002) *Biochem. Biophys. Res. Commun.* **291,** 1113–1118
16. Saxena, I. M., Brown, R. M., Jr., and Dandekar, T. (2001) *Phytochemistry* **57,** 1135–1148
17. Wandall, H. H., Pedersen, J. W., Park, C., Levery, S. B., Pizette, S., Cohen, S. M., Schwientek, T., and Clausen, H. (2002) *J. Biol. Chem.* **278,** 1411–1414
18. Schwientek, T., Keck, B., Levery, S. B., Jensen, M. A., Pedersen, J. W., Wandall, H. H., Stroud, M., Cohen, S. M., Amado, M., and Clausen, H. (2002) *J. Biol. Chem.* **277,** 32421–32429
19. Goode, S., Melnick, M., Chou, T. B., and Perrimon, N. (1996) *Development* **122,** 3863–3879
20. Panin, V. M., and Irvine, K. D. (1998) *Semin. Cell Dev. Biol.* **9,** 609–617
21. Perrimon, N., Engstrom, L., and Mahowald, A. P. (1989) *Genetics* **121,** 333–352
22. Nishikawa, K., Matsuoka, K., Kita, E., Okabe, N., Mizuguchi, M., Hino, K., Miyazawa, S., Yamasaki, C., Aoki, J., Takashima, S., Yamakawa, Y., Nishijima, M., Terunuma, D., Kuzuhara, H., and Natori, Y. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7669–7674
23. de Maagd, R. A., Bravo, A., and Crickmore, N. (2001) *Trends Genet.* **17,** 193–199
24. Jurat-Fuentes, J. L., Gould, F. L., and Adang, M. J. (2002) *Appl. Environ. Microbiol.* **68,** 5711–5717
25. Kumaraswami, N. S., Maruyama, T., Kurabe, S., Kishimoto, T., Mitsui, T., and Hori, H. (2001) *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.*. **129,** 173–183
26. Carriere, Y., Ellers-Kirk, C., Sisterson, M., Antilla, L., Whitlow, M., Dennehy, T. J., and Tabashnik, B. E. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 1519–1523

the FEBS Journal

Characterization of mucin-type core-1 β1-3 galactosyltransferase homologous enzymes in *Drosophila melanogaster*

Reto Müller[1], Andreas J Hülsmeier[1], Friedrich Altmann[2], Kelly Ten Hagen[3], Michael Tiemeyer[4] and Thierry Hennet[1]

1 Institute of Physiology, University of Zürich, Switzerland
2 Institute of Chemistry, Universität für Bodenkultur, Wien, Austria
3 Developmental Glycobiology Unit, National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD, USA
4 Complex Carbohydrate Research Center, The University of Georgia, Athens, GA, USA

Keywords
Drosophila; galactosyltransferase; glycolipid; glycosylation; mucin

Correspondence
T. Hennet, University of Zürich, Institute of Physiology, Winterthurerstrasse 190, 8057 Zürich, Switzerland
Fax: +41 44 6356814
E-mail: thennet@access.unizh.ch

(Received 22 March 2005, revised 11 June 2005, accepted 29 June 2005)

doi:10.1111/j.1742-4658.2005.04838.x

Mucin type O-glycosylation is a widespread modification of eukaryotic proteins. The transfer of *N*-acetylgalactosamine to selected serine or threonine residues is catalyzed by a family of polypeptide *N*-acetylgalactosaminyltransferases localized in the Golgi apparatus. The most abundant elongation of O-glycans is the addition of a β1-3 linked galactose by the core-1 β1-3 galactosyltransferase (core-1 β3GalT), thereby building the T-antigen or core-1 structure Gal(β1-3)GalNAc(α1-O). We have isolated four *Drosophila melanogaster* cDNAs encoding proteins structurally similar to the human core-1 β3GalT enzyme and expressed them as FLAG-tagged proteins in Sf9 insect cells. The identity of these *D. melanogaster* β3GalT enzymes with a core-1 β3GalT activity was confirmed by utilization of MUC5AC mucin derived O-glycopeptide acceptors. In addition to the core-1 β3GalT activity toward O-glycoprotein substrates, one member of this enzyme family showed a strong activity towards glycolipid acceptors, thereby building the core-1 terminated Nz6 glycosphingolipid. Transcripts of the embryonically expressed core-1 β3GalTs were found in the maternally deposited mRNA, in salivary glands and in the amnioserosa. The presence of multiple core-1 β3GalT genes in *D. melanogaster* suggests an increased complexity of core-1 O-glycan expression, which is possibly related to multiple developmental and physiological functions attributable to this class of glycans.

Mucin-type O-glycosylation is initiated by the transfer of *N*-acetylgalactosamine (GalNAc) to the hydroxyl group of selected serine and threonine residues. This transfer is catalyzed by a family of polypeptide *N*-acetylgalactosaminyltransferase (ppGalNAcT) enzymes localized in the Golgi apparatus [1]. The resulting GalNAc(α1-O)Ser/Thr epitope, also known as the Tn-antigen [2], is elongated in most cells by the addition of galactose (Gal) via a β1-3 linkage, thus forming the core-1 Gal(β1-3)GalNAc(α1-O) structure. Whereas more than 15 ppGalNAcTs have been identified in mammalian genomes, only a single core-1 β1-3 galactosyltransferase (β3GalT) enzyme has been described to date [3,4]. The importance of the early core-1 β3GalT

Abbreviations
2AB, 2-aminobenzamide; DIG, digoxigenin; β3GalT, β1-3 galactosyltransferase; Gal, galactose; GalNAc, *N*-acetylgalactosamine; GU, glucose unit; ppGalNAcT, polypeptide *N*-acetylgalactosaminyltransferase; TBS, Tris-buffered saline.

activity was demonstrated by the embryonic lethality observed in mice bearing an inactivated *core-1* β3GalT gene [5]. These core-1 β3GalT1-null mice exhibited angiogenesis defects and hemorrhages possibly caused by defective interactions between endothelial cells and the extracellular matrix, highlighting the significance of core-1 mucin structures in mammalian development.

Nine ppGalNAcT genes have been described in *D. melanogaster* [6] but no *core-1* β3GalT gene has been characterized up to now. As shown by peanut agglutinin binding, the distribution of core-1 glycans is regulated in a tissue- and stage-specific manner during embryonic development in *D. melanogaster* [7,8]. Core-1 glycans are found on mucin glycoproteins isolated from different *D. melanogaster* cell lines and tissues [9–11]. In addition, core-1 glycans occur on short antibacterial peptides such as Drosocin in *Drosophila* [12] and Diptericin in *Phormia* [13]. Remarkably, the O-glycan moiety of these peptides increases their antibacterial activity.

Protein sequence domains of glycosyltransferases are typically conserved between animal species, thus facilitating the identification of orthologous proteins across genomes. However, structural similarity alone is insufficient to conclusively assign an enzymatic activity to a novel protein as structurally related proteins may actually utilize different acceptor and donor substrates. To better understand the molecular pathways of O-glycosylation in insects, we have isolated the four closest homologous cDNAs to the human core-1 β3GalT in *D. melanogaster* and characterized their respective enzymatic activity and their expression pattern during early development.

Results

A search for *D. melanogaster* genes encoding proteins similar to the mammalian core-1 β3GalT enzymes yielded several hits as noted previously [3]. Using the TBLASTN algorithm [14] on the *D. melanogaster* genome sequence available through the BDGP server (http://www.fruitfly.org), we retrieved the cDNAs encoding the four closest homologous proteins to the human *core-1* β3GalT enzyme (Fig. 1). The overall sequence identity ranged from 31% to 43%, whereas several regions were highly conserved between the retrieved proteins and the human core-1 β3GalT. The detection of conserved TWG, DDD and EDV motifs, which are typical of β1,3 glycosyltransferase proteins [15], supported the potential functional orthology with the core-1 β3GalT enzyme (Fig. 1). The amino acid sequences retrieved from the *D. melanogaster* genome

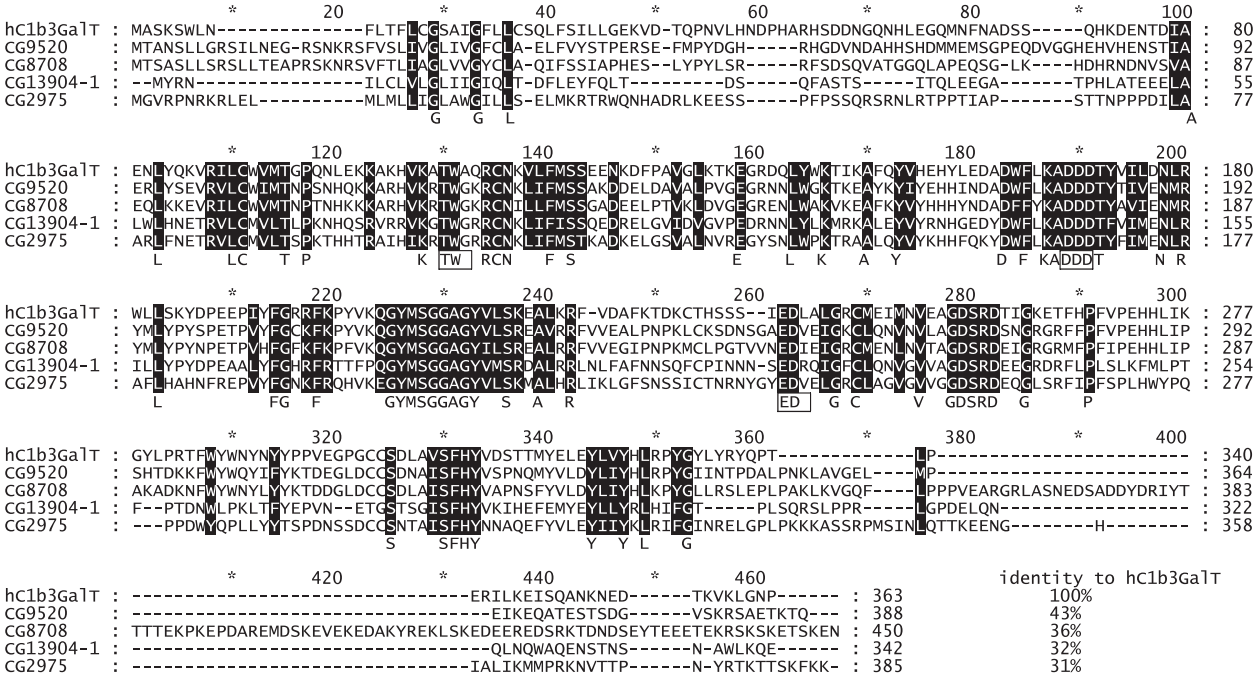

Fig. 1. Alignment of core-1 β3GalT candidate proteins. CLUSTALW [34] alignment of the human core-1 β3GalT protein (hC1b3GalT, accession: NP_064541) and of four similar *D. melanogaster* proteins. Amino acids conserved in all proteins are shaded in black. The TWG-, DXD- and EDV-motifs are boxed. Percentages of sequence identity of the *D. melanogaster* proteins to the human *core-1* β3GalT are indicated in the final column.

were in agreement with the gene models proposed by Flybase, with the exception of CG13904-1. The candidate protein of Flybase, i.e., CG13904, was modeled as a fusion protein by the gene prediction algorithm, where it represents a large protein of 680 amino acids with two similar domains. However, a comparison of this model with canonical β1-3 glycosyltransferases suggested that CG13904 represented two distinct genes arranged in tandem. However, we were unable to isolate a cDNA with an ORF consistent with a full length protein encoded by the 3′ located gene of the CG13904 locus either from adult, embryonic, or Schneider-2 cell cDNA. The cDNAs representing CG9520, CG8708 and CG13904-1 were isolated from embryonic mRNA whereas the cDNA for CG2975 could not be found in embryonic, but only in larval and adult mRNA.

The retrieved candidate cDNAs were expressed as N-terminally FLAG-tagged recombinant proteins in Sf9 cells. The expression of full-length recombinant proteins in Sf9 cells was confirmed by western blot analysis based on the detection of the FLAG-epitope (data not shown). The presence of this FLAG-epitope also enabled the capture and partial purification of the recombinant proteins for further characterization. To assay the enzymatic activity of each candidate protein, we first tested the transfer of Gal to GalNAc(α1-O)Bz using equal amounts of FLAG-recombinant proteins. CG9520 exhibited a high activity, whereas CG13904-1, CG2975 and CG8708 were only moderately active (Table 1). The screening for possible additional glycosyltransferase activity was extended by assaying the donor substrates UDP-Gal, UDP-GalNAc, UDP-GlcNAc, UDP-GlcA and UDP-Glc against the acceptor monosaccharides Gal, GalNAc, GlcNAc, Glc, fucose, mannose and xylose, each derivatized to pNP in either α and β anomeric configuration. We also tested various assay conditions with different detergents and detergent concentrations, by using other divalent cations, by applying a range of pH and temperature. The four enzymes showed similar requirement for Mn^{2+} and were most active at 25 °C, pH 6.6 and in the presence of 0.4% (v/v) Triton X-100. The enzymes were more active toward α-anomeric over β-anomeric monosaccharides with a marked preference for GalNAc(α1-O)Bz. CG9520 showed also a pronounced galactosyltransferase activity toward GlcNAc(α1-O)pNP, Gal(α1-O)pNP, GalNAc(β1-O)pNP and Man(α1-O)pNP (Table 1).

Table 1. Monosaccharide acceptor specificity of *D. melanogaster* core-1 β3GalT homologs.

	Enzyme[a] (pmol Gal·min^{-1}·mL^{-1})				
Acceptor (10 mM)	BRN[b]	CG9520	CG8708	CG13904-1	CG2975
GalNAc(α1-O)Bz	36	27 415	107	170	182
GalNAc(β1-O)pNP	30	1126	30	39	37
GlcNAc(α1-O)pNP	24	14 426	55	120	32
GlcNAc(β1-O)pNP	25	76	21	35	27
Gal(α1-O)pNP	27	2411	32	43	30
Gal(β1-O)pNP	38	44	30	37	34
Glc(α1-O)pNP	22	62	28	31	28
Man(α1-O)pNP	29	205	22	59	29
Fuc(α1-O)pNP	31	58	23	40	31
Xyl(α1-O)pNP	35	64	24	37	31

[a] Anti-FLAG-beads bound lysate of Sf9 cells. [b] The *D. melanogaster* β1-3 *N*-acetylglucosaminyltransferase *brainiac* (BRN) was used as negative control.

To verify that the active *D. melanogaster* core-1 β3GalT homologs indeed yielded a β1-3 linkage, we produced 10 nM of galactosylated GalNAc(α1-O)Bz using each of the four active galactosyltransferases CG9520, CG8708, CG13904-1 and CG2975 and analyzed their respective product by HPLC and MS. The disaccharides generated were first isolated by normal-phase chromatography. The product peaks were identified by electrospray-MS by their mass of 496.16 Da ([M + Na^+] ion). The linkage of the GalNAc residue in the disaccharide was investigated by permethylation analysis. In the gas-chromatographic separation of partially methylated alditol acetates, the GalNAc derivative eluted slightly after the derivative from a 4-substituted GlcNAc (reference made from bovine fetuin; 15.1 vs. 14.3 min). Partially methylated alditol acetates yield characteristic fragmentation patterns dependant on the substitution positions of a residue [16]. The GalNAc derivative gave fragment ions which strongly indicated a 3-substitution of the acceptor GalNAc whereas ions pointing at a 4- or 6-substitution were missing (Fig. 2).

Considering the artificial nature of the GalNAc(α1-O)Bz substrate, we also measured the core-1 β3GalT activity of the four active *D. melanogaster* enzymes towards various GalNAc(α1-O)glycopeptide, glycoprotein and glycolipid acceptors. The GalNAc(α1-O)glycopeptides assayed were derived from the MUC5AC sequence GTTPSPVPTTSTTSAP, where either Thr at position 3 (MUC5AC-3), Thr at position 13 (MUC5AC-13) or both Thr3 and Thr13 residues (MUC5AC-3/13) carried a GalNAc(α1-O) monosaccharide. These glycopeptides have been shown to act as substrates for mammalian and *D. melanogaster* ppGalNAcT enzymes [6]. Whereas CG9520 was able to transfer Gal to the three glycopeptides at equal efficiency, CG8708 showed a preference for the diglycosylated peptide MUC5AC-3/13 and CG13904-1 was more active toward MUC5AC-13 and MUC5AC-3/13

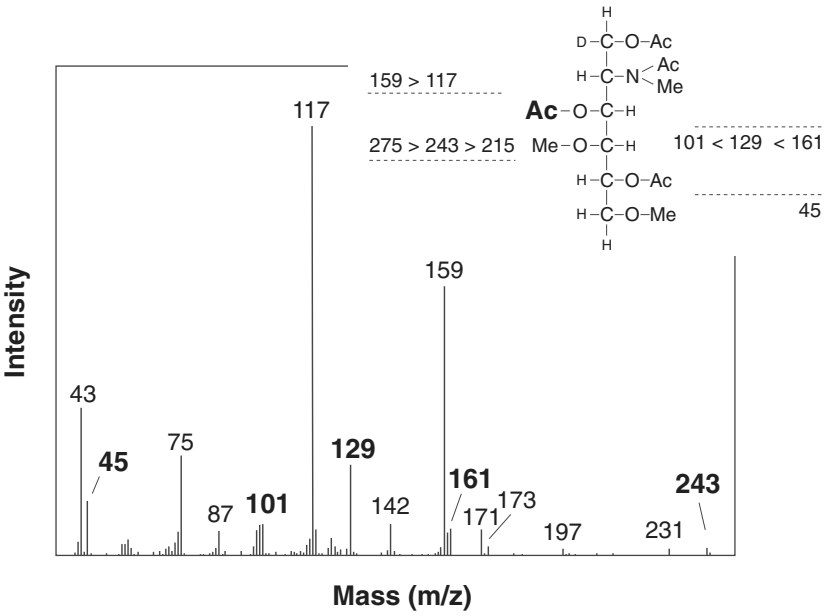

Fig. 2. Linkage analysis of the disaccharide Gal-GalNAc. The fragment spectrum of the partially methylated alditol acetate derived from the GalNAc residue is shown together with a fragmentation scheme. Diagnostic fragments are shown in bold. Equally important is the absence of fragments pointing at a 4- (e.g. 233 and 203) or 6-linkage (e.g. 189 and 203).

Table 2. Specificity of *D. melanogaster* core-1 β3GalT homologs toward complex type acceptors.

		Enzyme[a]				
Acceptor type	Name	BRN[b]	CG9520	CG8708	CG13904-1	CG2975
Glycopeptide	MUC5AC-3[c]	2	10 225	19	97	0
(pmol Gal·min^{-1}·mL^{-1})	MUC5AC-13[c]	0	12 398	184	184	0
	MUC5AC-3/13[c]	0	12 718	442	201	0
Glycoprotein	asOSM[d]	8	266	17	8	8
(pmol Gal·min^{-1}·mL^{-1})	asBSM[e]	3	223	4	3	5
Glycolipid (d.p.m.·h^{-1})	Sf9[f]	19	80	8	12	12
	Schneider-2[f]	12	1128	15	12	9

[a] Anti-FLAG-beads bound lysate of Sf9 cells. [b] The *D. melanogaster* β1-3 *N*-acetylglucosaminyltransferase *brainiac* (BRN) was used as negative control. [c] O-glycopeptide MUC5AC acceptors assayed at 2.5 mM (≈ 4.5 μg·μL^{-1}). Amino acids with GalNAc are in parentheses. MUC5AC-3, GT[T]PSPVPTTSTTSAP; MUC5AC-13, GTTPSPVPTTST[T]SAP; MUC5AC-3/13, GT[T]PSPVPTTST[T]SAP. [d] asOSM, asialo-ovine submaxillary mucin, assayed at 1.5 μg·μL^{-1}. [e] asBSM, asialo-bovine submaxillary mucin, assayed at 0.35 μg·μL^{-1}. [f] Assayed at 0.1 μg mannose equivalents·μL^{-1}.

(Table 2). CG2975 was inactive towards the three MUC5AC glycopeptides, although control reactions using GalNAc(α1-O)Bz confirmed the inherent galactosyltransferase activity of this protein. By comparison, when typical core-1 containing mucin glycoproteins were used as acceptors, only CG9520 showed a significant galactosyltransferase activity against asialo-ovine and asialo-bovine submaxillary mucins (Table 2).

Drosophila melanogaster glycolipids have been shown to contain the Gal(β1-3)GalNAc terminal epitope, as for example found in the Nz6 glycolipid Gal(β1-3)GalNAc(α1-4)GalNAc(β1-4)[phosphoethanolamine-6]GlcNAc(β1-3)Man(β1-4)Glc(β1-O)Cer [17,18]. To analyze whether *D. melanogaster* core-1 β3GalT homologs could catalyze the elongation of glycolipid substrates, we tested total glycolipids isolated from the *D. melanogaster* Schneider-2 cells and from *Spodoptera frugiperda* Sf9 cells as possible acceptors. Only CG9520 was able to transfer Gal to glycolipid acceptors, and this only to Schneider-2 derived glycolipids (Table 2). Considering this significant activity of CG9520 towards Schneider-2 glycolipids, we have analyzed the products of this reaction by TLC and HPLC. The TLC profile of *in vitro* [^{14}C]Gal-labeled Schneider-2 glycolipids showed several products, termed A–E in Fig. 3, which were isolated from the TLC and subjected to ceramide glycanase digestion. The released glycans were derivatized with 2-aminobenzamide (2AB) prior to GlycoSep–N normal

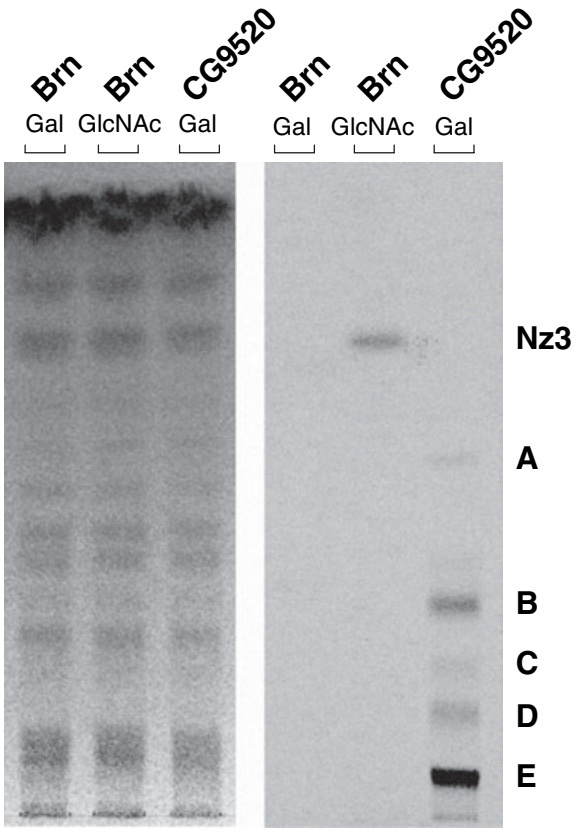

Fig. 3. Extension of glycolipids by CG9520. Glycolipids isolated from Schneider-2 cells were incubated with CG9520 and with the β1-3 *N*-acetylglucosaminyltransferase brainiac (BRN) together with the donor substrates indicated, i.e. UDP-[^{14}C]Gal or UDP-[^{14}C]GlcNAc. Reaction products were separated by TLC and detected by orcinol staining (left panel) and autoradiography for 24 h (right panel). The position of the BRN glycolipid product Nz3 [17] is marked in the right margin and the five products resulting from CG9520 extension are marked from A to E.

phase chromatography, calibrated with 2AB-labeled dextran oligomers to allow the expression of the retention times as glucose units (GU) (Fig. 4). Of the TLC bands analysed, the glycan released from band B co-eluted with authentic Nz6 saccharide [17] at 6.09 GU. The ceramide glycanase products released from bands A, C and D differed in their elution position of about one GU from the Nz6 saccharide (Fig. 4). The similar HPLC profiles obtained for C and D likely accounts for the loss of acid labile groups after mild acid hydrolysis treatment. The 2AB-glycan isolated from band E coeluted with authentic octaosylceramide Nz8 saccharide at 7.94 GU, suggesting that E could represent Gal-extended Nz7. This result underlined the function of CG9520 as a possible Nz6-synthesizing enzyme.

The patterns of *core-1* β3GalT gene expression were investigated during early fly development by *in situ* labeling in whole mount embryos with digoxigenin (DIG)-labeled probes. CG9520 mRNA was deposited into the embryo by the mother, was lost quickly thereafter and reappeared at around stage 9–10 to be expressed in a wide stripe in the amnioserosa of the embryo (Fig. 5), which is required for dorsal closure during fly development [19]. Finally, the staining followed the vanishing amnioserosa. By contrast, the two late embryonically expressed CG8708 and CG13904-1 genes were both expressed solely in salivary glands (Fig. 6).

Discussion

In the present study, we have shown that several *D. melanogaster* β3GalT enzymes can produce the mucin-type core-1 structure when assayed *in vitro*. The O-glycan core-1 biosynthetic activity could be established for three of these enzymes, as shown by the successful galactosylation of MUC5AC mucin derived glycopeptides. The comparison between the activity of *D. melanogaster* core-1 β3GalT enzymes towards MUC5AC glycopeptides showed a substrate preference associated with the glycopeptide structure itself because CG8708 preferred the diglycopeptide MUC5AC-3/13. The fact that these two core-1 β3GalT enzymes hardly glycosylated typical O-glycoproteins such as the asialo-ovine and asialo-bovine submaxillary mucins also speaks for a recognition of the peptide sequence itself by core-1 β3GalT proteins. In addition to O-glycopeptide acceptors, the CG9520 enzyme described here was able to transfer Gal to neutral glycolipids isolated from *D. melanogaster* Schneider-2 cells. The multiple reaction products identified after TLC and HPLC analysis showed that CG9520, considering its loose acceptor specificity (Table 1), probably added Gal to glycolipids of the Nz-series terminated with αGalNAc, βGalNAc and βGlcNAc such as Nz5, Nz4/Nz8 and Nz7, respectively [17]. The low core-1 β3GalT activity detected for CG8708 and CG13904-1 in comparison to that of CG9520 could indicate that they do not represent true core-1 β3GalT enzymes. However, as mentioned above, it is also possible to explain this difference if the enzymes do recognize the peptide backbone in the context of the acceptor substrate. Similarly, the characterization of the family of ppGalNAcT in several organisms has shown that the glycosyltransferase activities measured *in vitro* can vary over several orders of magnitude depending on the substates applied [6,20].

In mammalian cells, proper core-1 β3GalT activity has been shown to rely on interactions with the structurally related cosmc protein, which is devoid of glycosyltransferase activity but acts as a chaperone

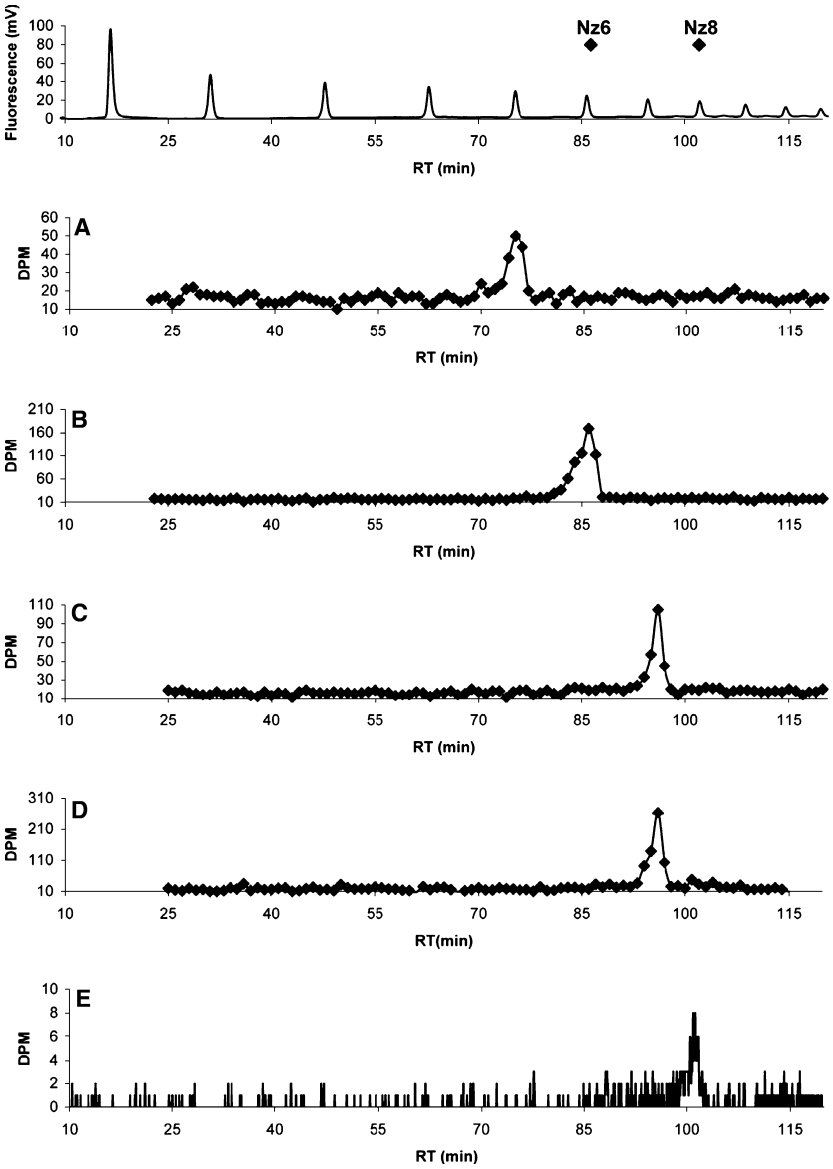

Fig. 4. HPLC profiling of glycolipid-derived oligosaccharides. The upper panel shows the normal phase chromatography fluorescence profile of 2AB labeled dextran oligomers corresponding to GU1-11. The elution positions of 2AB labelled Nz6 and Nz8 saccharides derived from authentic *D. melanogaster* glycolipids [17] are indicated by diamonds at 6.09 and 7.94 GU, respectively. (A–E) show the elution profiles of [^{14}C]Gal-labeled, ceramide glycanase released and 2AB-derivatized glycolipid saccharides isolated from the corresponding TLC bands A-E (see Fig. 3).

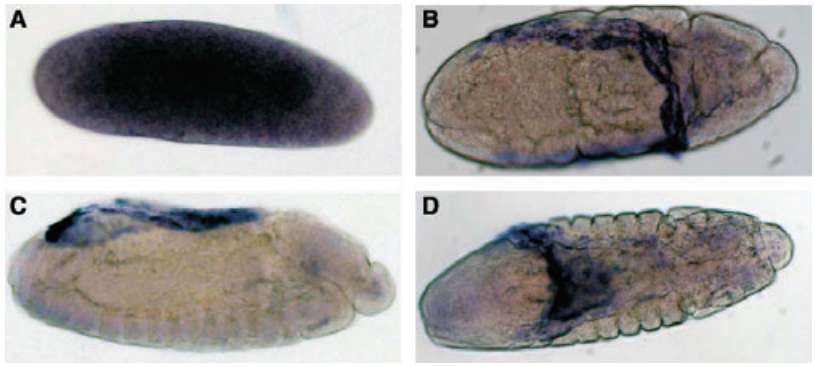

Fig. 5. Embryonic localization of CG9520 transcripts. The expression pattern of the CG9520 gene was detected by whole mount *in situ* hybridization during early *D. melanogaster* development. (A) Stage-2 embryo displaying the maternal deposition of CG9520 mRNA in the embryo. (B) Stage-11 embryo with staining in the amnioserosa. (C) Lateral view of a stage-12 embryo; (D) Dorsal view of stage-12 embryo.

for the core-1 β3GalT enzyme [21]. Whereas no homologous sequence to *cosmc* could be retrieved from the *D. melanogaster* genome, we did identify, in addition to the four *core-1* β3GalT cDNAs characterized here, five more genes showing a similarity to *core-1* β3GalT between 28 and 33% at the protein

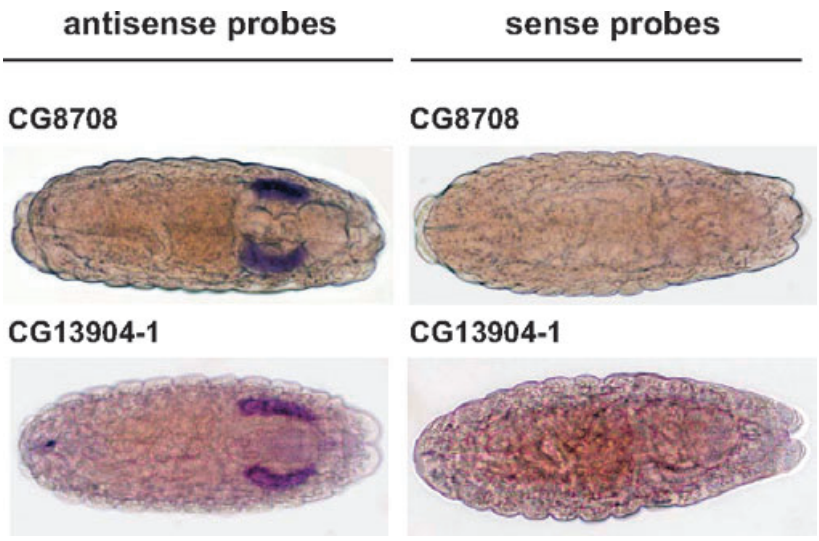

Fig. 6. Salivary gland expression of CG8708 and CG13904-1. The expression of the two *core-1* β3GalT genes during embryogenesis was confined to salivary glands. The four panels show ventral views of stage-16 embryos.

sequence level. The expression of these five genes in Sf9 cells failed to reveal any glycosyltransferase activity (data not shown), suggesting that some of these inactive proteins may act like cosmc as chaperones for core-1 β3GalT. However, the combined co-expression of active and inactive *D. melanogaster* core-1 β3GalT enzymes did not affect in any manner the glycosyltransferase activity measured in Sf9 cells (data not shown).

In the present study, we have reported the presence of at least three *core-1* β3GalT genes in the *D. melanogaster* genome. One reason for this higher number of core-1 β3GalTs in *D. melanogaster* may be related to differences in the regulation of gene expression between insects and mammals. The transcriptome of *D. melanogaster* is split into an adult and an embryonic one [22], potentially suggesting that the O-glycome of adult *D. melanogaster* may be constructed by glycosyltransferases that are not expressed during embryogenesis and early development. Alternatively, it is possible that insect core-1 β3GalT enzymes fulfil multiple tasks in various physiological processes. Adaptation to pathogens and to environmental stress often lead to lineage-specific expansion of gene clusters involved in such responses [23]. In this context, the specific expansion of *core-1* β3GalT genes in *D. melanogaster* may be interpreted in this way, as it has been observed for the lineage-specific expansion of glycosyltransferase families in animal genomes [24].

The expression patterns of the three embryonically expressed, active *core-1* β3GalT genes during early *D. melanogaster* development revealed the presence of transcripts in salivary glands and in the transient structure called amnioserosa. The presence of at least two ppGalNAcTs and two *core-1* β3GalTs suggests requirement of the T-antigen on proteins of the salivary glands. A potential target protein in embryonic salivary glands represents the secreted mucin-type glue protein encoded by the gene *salivary gland secretion 4* [25,26]. Salivary gland secrete is rich in carbohydrates and most salivary gland secreted proteins are suspected to be glycosylated because of their behavior in polyacrylamide gradients [26]. Previous studies based on lectin histochemistry with the Gal(β1-3)GalNAc-binding lectin peanut agglutinin failed to reveal any signal in embryonic salivary glands [8], which could mean that salivary O-glycan chains are elongated, thus abrogating peanut agglutinin binding. Furthermore, the peanut agglutinin staining in the developing nervous system documented by D'Amico and Jacobs [8] could not be confirmed in our *in situ* hybridization study. The comprehensive testing of all *core-1* β3GalT homologous genes during *Drosophila* development will show whether other genes are expressed in the tissues that are positive for peanut agglutinin binding.

Transcripts of the CG9520 *core-1* β3GalT gene were first detected as maternally deposited mRNA, in the amnioserosa and also in salivary glands. The amnioserosa separates two epithelial layers, the lateral and the dorsal epidermis until resorption of the yolk sac, allowing the epithelial layers to meet at the dorsal midline. The specific expression of CG9520 in the amnioserosa suggests a role for glycosylation in this process. However, the strong activity of the CG9520 enzyme towards glycolipid acceptors renders the interpretation of this potential involvement challenging. A precise structural analysis will be required to clarify whether O-glycoproteins or glycolipids mediate critical inter-

actions in the process of dorsal closure. In general, the dual acceptor specificity of CG9520 together with the identification of multiple *core-1* β3GalT enzymes in *D. melanogaster* will make it difficult to determine whether mucin-type O-glycosylation is essential for the development or survival of insects as it has been demonstrated for mammals using *core-1* β3GalT gene disruption in the mouse. However, the sophisticated genetics of the fruit fly as well as many available mutants should enable us to discern which of the members of this family are essential for development as well as eventually decipher their *in vivo* substrates.

Experimental procedures

Cloning of *Drosophila* cDNAs

Total RNA was extracted from tight-rod disintegrated 0–24 h embryo and adult *OregonR D. melanogaster* using Tri-Reagent (Sigma, St. Louis, MO, USA) according to the manufacturer's protocol. The isolated RNA (100 μg) was subjected to purification and mRNA selection using the GenElute^{TM} mRNA Miniprep Kit (Sigma). First strand cDNA was generated for 1 h at 37 °C using Omniscript reverse transcriptase (Qiagen, Hilden, Germany) primed with a polyT_{25} primer. The cDNAs of interest were amplified using specific primers and using the conditions listed in Table 3. The resulting fragments were subcloned into pBluescriptII SK$^+$ (Stratagene, La Jolla, CA, USA) and sequenced prior to transfer into pFastbac-FLAG vectors [27].

Expression of recombinant proteins

Recombinant baculoviruses containing the *D. melanogaster core-1* β3GalT candidate cDNAs were generated as described previously [28]. After infection of 1.5×10^7 *S. frugiperda* Sf9 insect cells with recombinant baculoviruses and incubation for 48 h at 27 °C, the cells were washed in 50 mM Tris-buffered saline (TBS), pH 7.4 and lyzed in 500 μL TBS containing 2% (v/v) Triton X-100, 10 μg·mL^{-1} benzamidine, 2 μg·mL^{-1} pepstatin A, 2 μg·mL^{-1} leupeptin, 2 μg·mL^{-1} antipain, 2 μg·mL^{-1} chymostatin and 0.2 mM phenylmethanesulfonyl fluoride (all from Fluka, Buchs, Switzerland). Post-nuclear supernatants were diluted to 1% (v/v) Triton X-100 in TBS and amounts of lysate corresponding to 5 mg total proteins were incubated with 120 μL EZview^{TM} Red Anti-FLAG-bead suspension (Sigma) under rotation for 10 h at 4 °C. Beads were washed three times with 2 mL ice-cold TBS and diluted to 25 μg total protein·μL^{-1} slurry. The integrity and amounts of FLAG-tagged recombinant proteins were inspected by western blotting.

Glycosyltransferase assays

Enzymatic activity towards p-nitrophenyl (pNP) and benzyl (Bz) derivatized monosaccharide acceptors (Sigma) was assayed using 250 μg bead-bound enzyme (10 μL) in 50 μL 100 mM cacodylate buffer pH 6.6, 20 mM $MnCl_2$, 5% (v/v) Me_2SO, 0.4% (v/v) Triton X-100, 0.2 μg·mL^{-1} 3×FLAG peptide (Sigma), 0.1 mM UDP-Gal (Fluka) including 2.5×10^4 c.p.m. UDP-[^{14}C]Gal (Amersham Biosciences, Arlington Heights, IL, USA), and 10 mM acceptor substrates (Table 1). Galactosyltransferase activity with CG9520 towards GalNAc(α1-O)Bz and GlcNAc(α1-O)pNP were measured with 0.5 mM UDP-Gal. Reactions were incubated at 25 °C for 10–30 min or overnight for acceptor screening, then stopped by incubation at 72 °C for 5 min. Reaction products were purified over C_{18} Sep-Pak cartridges (Waters, Milford, MA, USA) as described [28] and radioactivity was quantified in a Tri-Carb 2900TR liquid scintillation counter (Packard, Pangbourne, UK) with luminescence correction. Assays towards MUC5AC derived glycopeptide acceptors [6]

Table 3. Primers and conditions for molecular cloning of *D. melanogaster core-1* β3GalT homologs. Gene names are given according to Flybase (http://www.flybase.org) except for CG13904-1 (see main text). Restriction endonucleases used to clone PCR fragment into pBluescript SK+ are given in parenthesis and the corresponding restriction sites are underlined.

Gene		Annealing Temp (°C)	Fragment size (bp)
CG9520			
Forward	AAAACAAAAGCCAAATGACTGCCAAC (*Sma*I)	56.5	1188
Reverse	TGTCTAGATTATTGCGTCTTTGTCTCGGC (*Xba*I)		
CG8708			
Forward	AGGGATCCCACAATAAGTGCA GAATG (*Bam*HI)	56	1434
Reverse	GCGGTCTAGACTCAGAAACAG CTCAG (*Xba*I)		
CG2975			
Forward	GGAATTCCCTCAAGAGGAGCATAGAATG (*Eco*RI)	55.5	1232
Reverse	GCTCTAGAGCAGTCAATCCGAAATGAATG (*Xba*I)		
CG13904-1			
Forward	AGCTGGATCCGGTTAGTTGCAG (*Bam*HI)		
Reverse	TTGACTGTCGGTACCTTAAAATGAGTC (*Kpn*I)	57.5	1123

were carried out under similar conditions, except that the reaction volume was reduced to 25 μL, Me_2SO was omitted and using 0.1 mM UDP-Gal together with 5×10^4 c.p.m. UDP-[^{14}C]galactose. The enzymatic reaction was stopped by adding 500 μL cold H_2O. Samples were loaded on an AG1-X8 column (Bio-Rad, Hercules, CA, USA) and reaction products were eluted with H_2O. Assays towards the glycoprotein acceptors asialo-bovine submaxillary mucin (Sigma) and asialo-ovine submaxillary mucin (kindly provided by R.L. Hill, Duke University Medical Center, Durham, NC, USA) were carried out as described above for monosaccharide acceptor-based assays. Reaction products were precipitated with 1 mL cold 15% (v/v) trichloroacetic acid, 5% (v/v) phosphotungstic acid in H_2O, spotted on glass fiber filters (Whatman, Maidstone, UK) as described elsewhere [29] and measured in a scintillation β-counter.

Structural analysis

Dried mixtures containing GalNAc(α1-O)Bz and the product of the reaction with the galactosyltransferases studied were taken up in 80% (v/v) acetonitrile in water and subjected to normal phase HPLC on a TSKgel Amide-80 column (4.6 × 250 mm, Tosoh Bioscience, Montgomeryville, PA, USA) at a flow rate of 1 mL·min^{-1}. Solvent A was 50 mM ammonium formate at pH 4.4 and solvent B was 95% (v/v) acetonitrile. The column was equilibrated with 80% solvent B. After a 1-min hold postinjection the percentage of solvent B was lowered to 73%. Bz-glycosides were monitored at 254 nm. Peaks were examined by direct infusion electrospray-MS on a Q-Tof Global (Waters). Bz-disaccharide containing fractions were dried and permethylated using solid NaOH [30]. Partially permethylated alditol acetates were prepared using $NaBD_4$ as the reducing agent and analyzed by GC-MS using a 30 m/0.25 mm/0.25 μm HP5 column (Agilent, Palo Alto, CA, USA) and an Agilent GC-MS apparatus with helium as the carrier gas. Samples were injected with a low split at an oven temperature of 140 °C which was raised to 190 °C and to 260 °C with 10 and 4 °C·min^{-1}, respectively.

TLC

Glycolipids were extracted from *D. melanogaster* Schneider-S2 cells and 15 μg of mannose equivalents were used per glycosyltransferase assay as described previously [27] except that Triton X-100 was added to 1.4%. For TLC analysis, reaction products were dried under N_2, taken up in 100 μL H_2O and extracted 10 times with 900 μL toluene to remove Triton X-100 from the samples. Glycolipids were developed in chloroform/methanol/0.25% aqueous potassium chloride (10 : 10 : 3; v/v/v) on silica gel 60 aluminium high-performance TLC plates (Merck, Darmstadt, Germany). Plates were stained with orcinol sulfuric acid (Sigma) and autoradiographed for 24 h.

HPLC analysis

[^{14}C]Gal-labeled Schneider-S2 glycolipids (30 μg mannose equivalents) were developed by TLC and autoradiographed as outlined above. Radioactive bands were excised from the TLC plate and glycolipids were extracted from the silica matrix by sonication in methanol. Samples were subjected to mild acid hydrolysis in 40 mM trifluoroacetic acid in methanol/H_2O (1/1; v/v) for 10 min at 100 °C to eliminate acid labile glycan modifications [31], dried under N_2, taken up in 200 μL 50 mM sodium acetate pH 5.0, 0.75 mg·mL^{-1} sodium cholate (Sigma) prior to the addition of 0.2 U ceramide glycanase (Dextra Laboratory Ltd, Reading, UK) for a 24-h incubation at 37 °C, which was repeated for another 24 h. Reactions were stopped by extracting three times with 400 μL of H_2O-saturated butanol. The aqueous phase was dried briefly to remove residual butanol, subjected to a C_{18} Sep-Pak cartridge and ENVI-Carb column purification, 2AB derivatization and paper disk clean up as described [32] with minor modifications. Notably, samples were eluted from the ENVI-Carb column with 4 mL 50% (v/v) acetonitrile, subjected to 2AB-labeling and subsequent paper-disk clean up by placing the paper disk into 0.5 mL Ultrafree-MC filter devices (Millipore, Bedford, MA, USA). 2AB-labelled saccharides were eluted three times with 50 μL H_2O and aliquots were analyzed by GlycoSep–N normal phase chromatography [32] coupled to a Packard 500TR Series flow scintillation detector. Alternatively, 400-μL fractions were collected and radioactivity of each fraction was quantified with a Tri-Carb 2900TR liquid scintillation counter (Packard).

In situ hybridization

DIG-labeled RNA probes were prepared using the DIG RNA labeling Kit (Roche, Branchberg, NJ) by *in vitro* transcription with T7, T3 or SP6 RNA polymerase using pBluescript II SK$^+$ (Stratagene) or pGEM (Promega, Madison, WI, USA) derived DNA templates. Control reactions were carried out with sense transcripts. The probes, approximately 1 kb, were hydrolyzed for 90 min using standard procedures, precipitated with $LiCl_2$ and ethanol and quantified relative to each other following a protocol from the Berkley Drosophila Genome Project (BDGP) available at (http://www.bdgp.org/about/methods/Quantification_of_RNA.html). Equal amounts of DIG-labeled transcripts were used to probe 0–22-h-old y^1w^1 embryos following the method of Tautz and Pfeifle [33].

Acknowledgements

We thank Bea Berger and Marianne Farah for technical assistance. We also thank Dr Robert L. Hill for providing ovine submaxillary mucin. We are grateful to Drs

Eric Berger, Monika Hediger Niessen and Erich Frei for helpful suggestions and we acknowledge Dr Michael Gartner for making available the GC-MS equipment. This work was funded by the Swiss National Science Foundation Grant 631–062662.00 to T.H.

References

1 Ten Hagen KG, Fritz TA & Tabak LA (2003) All in the family: the UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases. *Glycobiology* **13**, 1R–16R.
2 Berger EG, Dinter A & Thurnher M (1994) Mucin type galactosyltransferase: enzymology and deficiency in the Tn syndrome. *Trends Glycosci Glycotechnol* **6**, 51–63.
3 Ju T, Brewer K, D'Souza A, Cummings RD & Canfield WM (2002) Cloning and expression of human core 1 beta 1,3 galactosyltransferase. *J Biol Chem* **277**, 178–186.
4 Ju T, Cummings RD & Canfield WM (2002) Purification, characterization and subunit structure of rat core 1 beta 1,3 galactosyltransferase. *J Biol Chem* **277**, 169–177.
5 Xia L, Ju T, Westmuckett A, An G, Ivanciu L, McDaniel JM, Lupu F, Cummings RD & McEver RP (2004) Defective angiogenesis and fatal embryonic hemorrhage in mice lacking core 1-derived O-glycans. *J Cell Biol* **164**, 451–459.
6 Ten Hagen KG, Tran DT, Gerken TA, Stein DS & Zhang Z (2003) Functional characterization and expression analysis of members of the UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase family from *Drosophila melanogaster*. *J Biol Chem* **278**, 35039–35048.
7 Fristrom DK & Fristrom JW (1982) Cell surface binding sites for peanut agglutinin in the differentiating eye disc of *Drosophila*. *Dev Biol* **92**, 418–427.
8 D'Amico P & Jacobs JR (1995) Lectin histochemistry of the *Drosophila* embryo. *Tissue Cell* **27**, 23–30.
9 Kramerov AA, Arbatsky NP, Rozovsky YM, Mikhaleva EA, Polesskaya OO, Gvozdev VA & Shibaev VN (1996) Mucin-type glycoprotein from *Drosophila melanogaster* embryonic cells: characterization of carbohydrate component. *FEBS Lett* **378**, 213–218.
10 Kramerova IA & Kramerov AA (1999) Mucinoprotein is a universal constituent of stable intercellular bridges in *Drosophila melanogaster* germ line and somatic cells. *Dev Dyn* **216**, 349–360.
11 Theopold U, Dorian C & Schmidt O (2001) Changes in glycosylation during *Drosophila* development. The influence of ecdysone on hemomucin isoforms. *Insect Biochem Mol Biol* **31**, 189–197.
12 Bulet P, Dimarcq JL, Hetru C, Lagueux M, Charlet M, Hegy G, Van Dorsselaer A & Hoffmann JA (1993) A novel inducible antibacterial peptide of Drosophila carries an O-glycosylated substitution. *J Biol Chem* **268**, 14893–14897.
13 Bulet P, Hegy G, Lambert J, van Dorsselaer A, Hoffmann JA & Hetru C (1995) Insect immunity. The inducible antibacterial peptide diptericin carries two O-glycans necessary for biological activity. *Biochemistry* **34**, 7394–7400.
14 Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
15 Malissard M, Dinter A, Berger EG & Hennet T (2002) Functional assignment of motifs conserved in β1,3-glycosyltransferases. *Eur J Biochem* **269**, 233–239.
16 Lindberg B & Lonngren J (1978) Methylation analysis of complex carbohydrates: general procedure and application for sequence analysis. *Methods Enzymol* **50**, 3–33.
17 Seppo A, Moreland M, Schweingruber H & Tiemeyer M (2000) Zwitterionic and acidic glycosphingolipids of the *Drosophiphila melanogaster* embryo. *Eur J Biochem* **267**, 3549–3558.
18 Wiegandt H (1992) Insect glycolipids. *Biochim Biophys Acta* **1123**, 117–126.
19 Jacinto A & Martin P (2001) Morphogenesis: unravelling the cell biology of hole closure. *Curr Biol* **11**, R705–R707.
20 Hagen FK & Nehrke K (1998) cDNA cloning and expression of a family of UDP-*N*-acetyl-*D*-galactosamine: polypeptide-*N*-acetylgalactosaminyltransferase sequence homologs from *Caenorhabditis elegans*. *J Biol Chem* **273**, 8268–8277.
21 Ju T & Cummings RD (2002) A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. *Proc Natl Acad Sci USA* **99**, 16613–16618.
22 Spellman PT & Rubin GM (2002) Evidence for large domains of similarly expressed genes in the *Drosophila* genome. *J Biol* **1**, 5.
23 Aravind L & Subramanian G (1999) Origin of multicellular eukaryotes – insights from proteome comparisons. *Curr Opin Genet Dev* **9**, 688–694.
24 Lespinet O, Wolf YI, Koonin EV & Aravind L (2002) The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Res* **12**, 1048–1059.
25 Barnett SW, Flynn K, Webster MK & Beckendorf SK (1990) Noncoordinate expression of *Drosophila* glue genes: Sgs-4 is expressed at many stages and in two different tissues. *Dev Biol* **140**, 362–373.
26 Beckendorf SK & Kafatos FC (1976) Differentiation in the salivary glands of *Drosophila melanogaster*: characterization of the glue proteins and their developmental appearance. *Cell* **9**, 365–373.
27 Müller R, Altmann F, Zhou D & Hennet T (2002) The *Drosophila melanogaster brainiac* protein is a

glycolipid-specific β1,3 N-acetylglucosaminyltransferase. *J Biol Chem* **277**, 32417–32420.

28 Hennet T, Dinter A, Kuhnert P, Mattu TS, Rudd PM & Berger EG (1998) Genomic cloning and expression of three murine UDP-galactose: β-*N*-acetylglucosamine β1,3-galactosyltransferase genes. *J Biol Chem* **273**, 58–65.

29 Malissard M, Borsig L, Di Marco S, Grutter MG, Kragl U, Wandrey C & Berger EG (1996) Recombinant soluble beta-1,4-galactosyltransferases expressed in *Saccharomyces cerevisiae*. Purification, characterization and comparison with human enzyme. *Eur J Biochem* **239**, 340–348.

30 Ciucanu I & Kerek F (1984) A simple and rapid method for the permethylation of carbohydrates. *Carbohydr Res* **131**, 209–217.

31 McConville MJ, Thomas-Oates JE, Ferguson MA & Homans SW (1990) Structure of the lipophosphoglycan from Leishmania major. *J Biol Chem* **265**, 19611–19623.

32 Grubenmann CE, Frank CG, Hülsmeier AJ, Schollen E, Matthijs G, Mayatepek E, Berger EG, Aebi M & Hennet T (2004) Deficiency of the first mannosylation step in the N-glycosylation pathway causes congenital disorder of glycosylation type Ik. *Hum Mol Genet* **13**, 535–542.

33 Tautz D & Pfeifle C (1989) A non-radioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* **98**, 81–85.

34 Thompson JD, Higgins DG & Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Conclusion and Outlook[*]

The determination of glycan structures and enzymological parameters of GTs remain essential tasks, but a systemic view on the function of the individual glycan is desirable. Therefore, the use of GT mutant alleles in the genetic analysis of glycan functions represents a step towards a better understanding of glycans. If no mutant for a GT locus is available, targeted overexpression (189), targeted disruption (190) or RNA interference mediated knock-down (191, 192) can be used to generate misexpressor lines, knock-out lines or phenocopies of null mutations. Structural differences between glycoconjugates from mutants and *wild-type* flies (98) also constitute the most convincing proof for the non-redundancy of an enzyme and the specificity for the glycan-biosynthetic pathway deduced from *in vitro* assays.

Along this line, it is largely unknown how the different glycosylation pathways are separated. For example, the linkage region of proteoglycans (Figure 13) displays the same terminal sequon as the Az sub-series of GSLs (Figure 15), GlcAβ1,3Galβ. However, GSLs do not seem to be a substrate for the highly processive GAG biosynthetic machinery *in vivo*, since no GlcA containing glycolipid has been found in dipterans which is elongated any further (82, 128, 193). Mechanisms for glycan-type separation have been exemplified in mammalian GSL biosynthesis: In some cells, lactosylceramide (Galβ1,4GlcβCer, c.f. Figure 16) synthase and the enzyme forming sialosyllactosylceramide were shown to form a complex (194), committing this GSL-species to the ganglioside pathway by the formation of a GT-complex. In other cells, lactosylceramide has been demonstrated to be sandwiched by a "glycolipid transfer protein" (195), committing it to the lacto-series pathway by a possible substrate protection mechanism. Genetic screens in *D. melanogaster* may also reveal genes responsible for the separation of individual glycan biosynthetic pathways.

To identify genes involved in glycan biosynthesis without prior knowledge of their sequence, a selection strategy involving toxic lectins as described with *C. elegans* (196, [II]) could be employed. For example, mutagenized *D. melanogaster* larvae would be fed specific *Bacillus thuringiensis* strains or toxins from other endospore forming bacilli. These toxins are used to control secondary pests of *D. melanogaster* in the vineyards of the Middle East (197). Some toxins that are active against *D. melanogaster*, encompassing the *Cry4* -(crystal toxin 4) family (personal communication Dr. Hala Khyami-Horani), display a lectin-like domain (198), indicating that they might require glycan binding for toxicity similarly to *Cry*14A or *Cry*5B in *C. elegans* [II]. While such selection strategies can limit time and expenses necessary to find

[*] Citations refer to section 11 of the introductory text to this thesis („Literature cited")

carbohydrate active enzymes expressed in the gut, visual inspection of phenotypes will still be required to score interactions of glycans in *D. melanogaster*. Genetic interaction experiments using GT loci allow addressing questions like: Which glycan interacts with which lectin? Which glycoconjugates require the glycan for function and which ones do not? Which signalling pathways are affected and at which level?

Since screening for interactors of GTs will represent a major organisational and technical challenge, one could resort to computational methods to identify a subset of candidate glycoconjugates and cross available mutants to GT alleles. Taking *core1* bearing proteins (Figure 19) as an example, neural network predictions could be used to look for O-GalNAc sites in a given amino acid sequence. At the moment, 76% of mammalian ppGalNAcT- substrates are correctly predicted (199). Furthermore, a metabolic labelling approach using the GalNAc analogue N-azidoacetylgalactosamine (GalNAz) was developed (200), eventually allowing large scale identification of mucin type O-glycosylated proteins in cell lines. On the other hand, lectins and other interactors will still have to be identified by screens, although biochemical methods as the use of glycan- arrays or –columns, and computational methods as sequence similarity searches with sequences of known lectins might identify candidates for glycan- binding proteins.

The genetic analysis of GT alleles is not without pitfalls and might, as in the case of HSPGs, display a plethora of interactions, which are difficult to disentangle because of interacting signalling pathways (discussed in ref. 201). Furthermore, GSLs have both the capability to bind lectins (16) and are structural components of *D. melanogaster* rafts (202). Rafts are two dimensional microdomains on membranes, which influence many different receptor ligand interactions (reviewed in ref. 203). For example, EGFR activation in mammals occurs upon removal of cholesterol from membranes (204), a method which destroys rafts. It may therefore be difficult to separate GSL-lectin interactions and raft-preserving functions of GSLs, a problem that also arises in the interpretation of the *brainiac* phenotype and may partially explain pleiotropic interactions. The case of *fringe*, which seems to primarily affect one receptor, may well be the exception. To circumvent these problems, it appears imperative to find the most isolated system to study mutant allele interactions, where the genetic information can be interpreted against the available background of biochemical data.

Appendix: Glycan biosynthetic activities in D. melanogaster[*]

Overview on the mucin type *O*-glycan synthesis pathway in *Drosophila melanogaster*.

Enzyme name (gene name)	Action of the transferase	Product	expressed in/during	Phenotypes	possible human orthologues[1] (by activity)
ppGalNAcTs					
CG8182 *pgant1*	Peptide/Glycopeptide transferase (160)	GalNAcα-peptide	all stages (RT-PCR)	n.d.	n.d.
CG3254 *pgant2*	Peptide/Glycopeptide Transferase (160)	GalNAcα-peptide	all stages (RT-PCR)	n.d.	ppGaNTase-T2
CG4445 *pgant3*	Peptide/Glycopeptide transferase (160, 162)	GalNAcα-peptide	all stages (RT-PCR)	n.d.	n.d.
CG31956 *pgant4*	Glycopeptide transferase (160)	GalNAcα-peptide	all stages (RT-PCR)	n.d.	n.d.
CG31651 *pgant5*	Peptide/Glycopeptide transferase (160)	GalNAcα-peptide	embryonic salivary glands, follicle cells (160) all stages (RT-PCR)	n.d.	ppGaNTase-T1/-T13
CG2103 *pgant6*	Glycopeptide transferase (160)	GalNAcα-peptide	embryonic salivary glands, follicle cells (160), all stages (RT-PCR)	n.d.	n.d.
CG6394 *pgant7*	Glycopeptide Transferase (160)	GalNAcα-peptide	all stages (RT-PCR)	n.d.	ppGaNTase-T7
CG7297 *pgant8*	Peptide/Glycopeptide transferase (160)	GalNAcα-peptide	all stages (RT-PCR)	n.d.	n.d.
CG7480 *pgant35*	Peptide/Glycopeptide transferase (160, 161, 178)	GalNAcα-peptide	all stages (RT-PCR)	late embryonic lethal	ppGaNTase-T11

[1] orthologous enzymes are inferred from the comparative ability to glycosylate a specific amino acid sequence

[*] Citations refer to section 11 of the introductory text to this thesis („Literature cited")

Enzyme name (gene name)	Action of the transferase	Product	expressed in/during	Phenotypes	possible human orthologues (by activity)
core1 β1,3GalTs					
CG9520	β1,3 GalT[III]	Galβ1,3GalNAcα -protein, GSLs (Nz6?)	maternal, amnioserosa [III]	n.d.	_core1_β1,3 GalT
CG8708	β1,3 GalT [III]	Galβ1,3GalNAcα -protein	embryonic and adult salivary glands [III], anterior ejaculatory duct (205)	n.d.	_core1_β1,3 GalT
CG13904-1	β1,3 GalTs [III]	Galβ1,3GalNAcα -protein,	salivary glands [III]	n.d.	_core1_β1,3 GalT
CG2975	β1,3 GalT [III]	Galβ1,3GalNAcα -aglycone?[3]	not expressed during embryogenesis [III]	n.d.	n.d.

Overview on the glycosphingolipid synthesis in *Drosophila melanogaster*.

Enzyme name (gene name)	Action of the transferase	Product	interactors/ pathway	Phenotypes	Orthologues
GlcT-1	β1,1 GlcT (206)	Glcβ1,1Cer	ceramide-induced apoptosis (206)	embryonic lethal (plasmid mediated RNAi) (206)	n.d.
Mactosylceramide synthase (*egghead*)	β1,4 ManT (123)	Manβ 1,4Glcβ 1,1Cer	TNFα, EGFR	late embryonic lethal, FDA, neurogenic, rough eyes (71, 136)	*C. elegans bre-3* no orthologues in mammals [II]
N3 synthase (*brainiac*)	β1,3 GlcNAcT ([I], (124)	GlcNAcβ1,3Manβ 1,4Glcβ 1,1Cer	TNFα, EGFR	late embryonic lethal, FDA, neurogenic (71, 137, 138)	*C. elegans bre*-5 no orthologues in mammals ([I], [II])
Nz3 synthase	β6 pEtnT	pEtnβ6-GlcNAcβ 1,3Manβ1,4Glcβ1,1Cer	n.d.	n.d.	n.d.
Nz4 synthase[1] CG8536?/ CG14517?	β1,4 GalNAcT (125, 207)	GalNAcβ1,4(pEtnβ6)GlcNAcβ1,3Manβ 1,4Glcβ1,1Cer	n.d.	n.d.	*C. elegans bre-4* [II]
Nz5 synthase CG17223	α1,4 GalNAcT (126)	GalNAcα 1,4GalNAcβ 1,4(pEtnβ6) GlcNAc β1,3Manβ1,4Glcβ 1,1Cer	n.d.	n.d.	n.d.

[1]CG8536 and CG14517 have been shown to catalyze the formation of the typical disaccharide, but have not been unambiguously assigned to theGSL-biosynthetic pathway. n.d.: not determined. FDA: fused dorsal appendage, *bre*: bacillus thuringiensis resistance c.f.www.wormbase.org

Enzyme name (gene name)	Action of the transferase	Product	interactors/ pathway	Phenotypes	Orthologues
Nz6 synthase CG9520[1]	β1,3 GalT [III]	Galβ1,3GalNAcα1,4GalNAcβ1,4(pEtnβ6)GlcNAcβ1,3Manβ1,4Glcβ1,1Cer	n.d.	n.d.	n.d.
Nz7 synthase	β1,3 GlcNAcT	GlcNAcβ 1,3Galβ 1,3GalNAcα1,4GalNAcβ1,4 (pEtnβ6) GlcNAcβ1,3Manβ1,4Glcβ1,1Cer	n.d.	n.d.	n.d.
$Nz_2$7 synthase	β6 pEtnT	pEtnβ6-GlcNAcβ 1,3Galβ1,3GalNAcα 1,4GalNAc β1,4(pEtnβ6)GlcNAcβ1,3Manβ 1,4Glcβ1,1 Cer	n.d.	n.d.	n.d.
$Nz_2$8 synthase[1] (CG8536? CG14517?)	β1,4 GalNAcT (125)	GalNAcβ 1,4(pEtnβ6)GlcNAcβ 1,3Galβ1,3 GalNAcα1,4GalNAcβ1,4(pEtnβ6)GlcNAcβ1,3 Manβ1,4Glcβ1,1 Cer	n.d.	n.d.	n.d.
$Nz_2$9 synthase	β1,3 GalT	Galβ1,3GalNAcβ 1,4 (pEtnβ6) GlcNAcβ1,3 GalNAcα1,4GalNAcβ1,4(pEtnβ6)GlcNAcβ 1,3 Manβ1,4Glcβ1,1Cer	n.d.	n.d.	n.d.
Az6 synthase[1] (CG32775? CG3881? CG6207?)	β1,3 GlcAT (127)	GlcAβ1,3Galβ1,3 GlcNAcβ1,3GalNAcα1,4 GalNAcβ1,4(pEtnβ6) GlcNAcβ1,3Man β1,4Glcβ1,1Cer	n.d.	n.d.	n.d.
$Az_2$9 synthase[1] (CG32775? CG3881? CG6207?)	β1,3 GlcAT (127)	GlcAβ 1,3Galβ 1,3GalNAcβ1,4(pEtnβ6)GlcNAcβ 1,3GalNAcα1,4GalNAcβ1,4(pEtnβ6)GlcNAcβ1,3Manβ1,4Glcβ1,1Cer	n.d.	n.d.	n.d.

n.d.: not determined, [1]CG9520, CG8536, CG14517, CG32775, CG3881 and CG6207 have been shown to catalyze the formation of the typical disaccharide, but have not been unambiguously assigned to the GSL-biosynthetic pathway. n.d.: not determined

.Overview on proteoglycan synthesis in *Drosophila melanogaster*

Gene	Action of gene product	Product, reaction	interactors/pathways affected	Phenotypes	Orthologues (*C. elegans*)
sugar-nucleotide synthesis					
sugarless (suppenkaspar, kiwi)	UDP-Glucose-6-dehydrogenase (94), EC 1.1.1.22	UDP-Glc => UDP-GlcA substrate for GlcATs	*wingless* (99), 118) Fibroblast growth factor pathways *(breathless, heartless* (208)	Segment polarity defects tracheal development disturbed	*sqv-4* (68)
CG7979?	UDP-GlcA decarboxylase EC 4.1.1.35	UDP-GlcA =>UDP-Xyl substrate for O-XylT	n.d.	n.d.	*sqv-1* (68)
sugar-nucleotide transporters					
fringe connection	UDP-GlcA,-GlcNAc, -Xyl Transporter (73, 117)	transport into Golgi Apparatus substrate for glycosyltransferases	*Notch* *wingless*	*fringe*-like phenotypes Segment polarity defects	*sqv-7* (68, 69)
slalom	adenosine 3'-phosphate 5'-phosphosulfate transporter (PAPS transporter) (209)	transport into Golgi Apparatus Substrate for sulfotransferases	*wingless*, *hedgehog* other pathways?	extra margin bristles in the wing, reduced Cubitus interruptus levels in wing-discs, ectopic wing-tissue around wing-vein L2 (209)	n.d.

Question mark: activity assignment solely by sequence similarity to known enzymes, n.d. not determined, *sqv*: *squashed vulva*, c.f.www.wormbase.org

Gene	Action of gene product	Product, reaction	interactors/ pathways affected	Phenotypes	Orthologues (*C. elegans*)
Glycosyltransferases involved in proteoglycan core tetra-saccharide synthesis					
o-xylosyltransferase OXT	UDP-Xylose: protein-β O-Xylosyl-Transferase (96) EC 2.4.2.26	Xylβ1,1O-Ser	n.d.	n.d.	*sqv-6*(68)
CG11780 β1,4 GalT-7	UDP-Gal: βXyl β1,4 GalT-I (210) EC 2.4.1.133	Galβ1,4 Xylβ1,1O-Ser	n.d.	Plasmid mediated RNA interference: late pupal lethal (210)	*sqv-3*(68, 69)
CG8734?	UDP-Gal: β1,3Gal β1,3 GalT-II EC 2.4.1.179	Galβ1,3Galβ1,4Xylβ1,4-O-Ser	n.d.	n.d.	*sqv-2*(68)
[1]CG32775? CG3881? CG6207?	UDP-GlcA: β1,3 Gal β1,3 GlcAT-I (127) EC 2.4.1.13	GlcAβ1,3 Galβ1,4Xylβ1,4-O-Ser	n.d.	n.d.	*sqv-8* (68, 69)

Question mark: activity assignment solely by sequence similarity to known enzymes (CG8734), [1]CG32775, CG3881 and CG6207 have been shown to catalyze the formation of the typical disaccharide, but have not been unambiguously assigned to the GAG-biosynthetic pathway. *sqv*: *squashed vulva*, c.f.www.wormbase.org

Gene	Action of gene product	Product, reaction	interactors/ pathways affected	Phenotypes	Orthologues (*C. elegans*)
Glycosyltransferases involved in Heparan Sulfate backbone synthesis					
brother of tout-velu[1]	UDP-GlcNAc: β1,3GlcA α1,4 GlcNAcT (Initiating GlcNAcT[1]) (97, 115) EC 2.4.1.17	GlcNAcα1,4 GlcAβ1,3 Galβ1,4Xylβ1,4-O-Ser	*wingless*, *decapentraplegic*, *hedgehog*	Trans-heterozygotes of the Ext family die as late pupae, Segment polarity defects (98). small eyes and legs (115)	*rib-2* (211)[1]
sister of tout-velu	GlcNAcα1,4GlcAβ1,3GlcA co-polymerase (97, 98, 115) EC 2.4.1.225	GlcAβ1,3GlcNAcα1,4 GlcAβ (activity as a complex with *tout-velu* (115))	Wingless - gradient disturbed, *wingless* target genes underexpressed (98), *hedgehog* (212-214)	trans heterozygotes of the Ext family die as late pupae, segment polarity phenotypes, *wingless*-like phenotypes in the wing, (98) small eyes and legs (115)	*rib-1*?
tout velu[1]	UDP-GlcA: GlcAβ3 GlcNAcα1,4 and UDP-GlcNAc: GlcNAcα1,4GlcA (co-polymerase) (94) EC 2.4.1.225	GlcAβ1,3GlcNAcα1,4 GlcAβ (98), (activity as a complex with *sister of tout-velu* (115))	*wingless*, d*ecapentraplegic*, *hedgehog* (98)	as above	*rib-2* (211)

[1] in contrary to *D. melanogaster tout-velu*, the *C. elegans* orthologue of *tout-velu* and *brother of tout-velu* has also chain initiating properties as shown in (211). *rib*: family of genes related to mammalian RIB-affecting EXT gene family, c.f.www.wormbase.org

Gene	Action of gene product	Product, reaction	interactors/ pathways affected	Phenotypes	Orthologues (*C. elegans*)
Heparan sulfate sulfotransferases and other modifiers					
sulfateless	heparin/heparan sulfate GlcNAc N-deacetylase/GlcN N-sulfotransferase	1. GlcNAc => GlcN GlcN =>N-sulfo-GlcN (94)	*wingless*, d*ecapentraplegic*, h*edgehog, dally* *frizzled* 2, FGF receptors (*breathless*, *heartless*) (107, 208) *dally like protein* (114)	Segment polarity phenotypes stomatogastric nervous system and second midgut constriction affected (107) Larval and pupal lethal (107)	n.d.
n.d.	heparin/heparan sulfate-glucuronic acid C5-epimerase	GlcA => IdoA	n.d.	n.d.	n.d.
pipe	heparin/heparan sulfate 2-O-sulfotransferase?	IdoA=>2-sulfo-IdoA?	Toll/Spaetzle Specification of Dorsal-Ventral axis of embryo	maternal effect dorsalized (215) overexpression: ventralization (216)	n.d.
hs6st	heparin/heparan sulfate 6-O-sulfotransferase (217)	GlcNac=>6-O-sulfo-GlcNac	FGF receptors (*breathless*)	Tracheal branch formation disturbed. Lethal (stage 17), (217)	n.d.
hs3st	heparin/heparan sulfate 3-O.sulfotransferase (218)	GlcNAc=>3-O-sulfo-GlcNAc	Notch	neurogenic (plasmid mediated RNAi (218))	n.d

Question mark: activity is an assumption based on sequence similarities to known enzymes

Gene	Action of gene product	Product, reaction	interactors/ pathways affected	Phenotypes	Orthologues (*C. elegans*)
Glycosyltransferases involved in chondroitin sulfate synthesis					
CG12913?	β1,4 GalNAcT-I	GalNAcβ1,4GlcAβ1,3 Galβ1,4Xylβ1,4-O-Ser	n.d.	n.d.	*sqv-5*
CG9220?	CS GlcAT-II or CS GlcAT-II co-polymerizing factor	GlcAβ1,3 GalNAcα GlcAβ1,3 Galβ1,4Xylβ1,4-O-Ser	n.d.	n.d.	*sqv-5*
CG4351?	CS GlcAT-II or CS GlcAT-II co-polymerizing factor	GlcAβ1,3 GalNAcα1,4 GlcAβ1,3 Galβ1,4Xylβ1,4-O-Ser	n.d.	n.d.	*sqv-5*

Question mark: assigned on the basis of sequence similarity to *C. elegans* and human enzymes solely. *sqv*: *squashed vulvae* (c.f. www.wormbase.org)

Curriculum vitae

Name	**Reto Müller**
Date and place of birth:	13/09/1973 in Grabs SG, CH
Marital status	single
Nationality:	Swiss
Home address:	Brunnenstrasse 2, 9470 Buchs

Education:

Oct 1988 to July 1993	*B-Matura* (Sargans, CH)
Nov 1993 to Oct1995	*Vordiplom* in *Biologie* II (University of Basel)
Sep 1995 to Sep 1998	Preparation of a three year Bioengineer's degree in Biotechnology at the *Ecole Superieure de Biotechnologie de Strasbourg* (ESBS*)*. European School of the Higher Rhine Universities, i.e. Basel, Strasbourg, Freiburg and Karlsruhe). Trinational diploma: *DEA biologie moleculaire et cellulaire* and *Ingenieur en biotechnologie* (France), *dipl. Biotech.* (Switzerland), *Diplom-Biotechnologe* (Germany).
Nov 1999 – present	Ph.D. student at the Institute of Physiology, Characterization of β1,3 Glycosyltransferases from *Drosophila melanogaster.*

Training periods:

1996 Jun and Jul	NMR spectroscopy, Biocomputing („Determination of the 3D structure of „enkelytin" using ROESY" at ESBS, University *Louis Pasteur*, Strasbourg, FR)
1996 Aug	fermentation, microbiology ("Esterases from *Sulfolobus acidocaldarius*", at the Institute of Botany,University of Zurich, CH)
1997 Jun	HPLC, protein purification (Large-scale preparation of the scorpio anti-microbial peptide androctonin" at *Institut de Biologie Moleculaire et Cellulaire*, Strasbourg, FR)
1997 Jul and Aug	GMP fermentation and assay development ("Esterases from *Sulfolobus acidocaldarius*" at *Fluka*, *Sigma-Aldrich*, Buchs SG, CH)
1998 Mar to Sep	Diploma thesis: "Detection of Rhamnagalacturonase II in various fungi" at *Gist-brocades* (now DSM), Delft, NL.
1999 Mar to Oct	Recombinant protein production at *Novartis*, Basel, CH

Publications:

** first author
* collaboration

**The *Drosophila melanogaster brainiac* protein is a glycolipid-specific beta 1,3N-acetylglucosaminyltransferase.
J Biol Chem. 2002 Sep 6;277(36):32417-20.

*Resistance to a bacterial toxin is mediated by removal of a conserved glycosylation pathway required for toxin-host interactions.
J Biol Chem. 2003 Nov 14;278(46):45594-602.

**Characterization of mucin-type core-1 beta1-3 galactosyltransferase homologous enzymes in Drosophila melanogaster.
Febs J 272, 4295-4305.